

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2001, 11:20:36 ; Search time 16.02 Seconds  
(without alignments)  
1426.490 Million cell updates/sec

Title: US-09-744-527-4  
Perfect score: 1566  
Sequence: 1 MEKTOETVQRIILEPYKYL.....NPEIYALVKHLKMKKEENE 300

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	797.5	50.9	428	2	S15662
2	563	37.2	335	2	S60921
3	310	19.8	347	2	A70139
4	292	18.6	327	2	D64407
5	257	16.4	364	2	C71295
6	238.5	15.2	322	2	C83075
7	233.5	14.9	325	2	T45152
8	231.5	14.8	299	2	C72242
9	231.5	14.8	325	2	G69165
10	231	14.8	304	2	A70377
11	217	13.9	323	2	A40433
12	217	13.9	323	2	T06874
13	213	13.6	342	2	E71101
14	210.5	13.4	323	2	E85982
15	209.5	13.4	324	2	A81213
16	209.5	13.4	322	2	S75427
17	208.5	13.3	323	2	E65109
18	207	13.2	330	2	A54058
19	201.5	12.9	346	2	C51539
20	200	12.8	327	2	D71651
21	199	12.7	297	2	E81400
22	197	12.6	332	2	A82688
23	195.5	12.5	347	2	H84270
24	195	12.5	348	2	H84630
25	192.5	12.3	300	2	F84365
26	192.5	12.3	326	2	F69535
27	189.5	12.1	390	2	F96813
28	189	12.1	272	2	G72410
29	185	11.8	319	2	F86792

30	182.5	11.7	329	2	I64160	hypothetical prote
31	182	11.6	348	2	A82322	octaprenyl-diphosp
32	180	11.5	335	2	E70549	probable heptapren
33	179.5	11.5	329	2	C75400	geranylgeranyl dip
34	178.5	11.4	379	2	C86306	prenyl transferase
35	177	11.3	337	2	B72560	probable geranylitr
36	175.5	11.2	291	2	D82778	geranyltransferasf
37	165	10.5	323	2	S76966	geranylgeranyl pyr
38	163.5	10.4	295	2	C64123	geranyltransferasf
39	163.5	10.4	323	2	S73189	prenyl transferase
40	163	10.4	320	2	B73457	polyprenyl synthas
41	159	10.2	272	2	A69961	geranyltransferasf
42	158.5	10.1	326	2	E86647	prenyl transferase
43	158	10.1	300	2	T06969	farneisyltransstrans
44	157.5	10.1	297	2	A84759	probable trans-pre
45	154.5	9.9	291	2	E71491	probable geranyl t

## ALIGNMENTS

RESULT 1  
S15662  
farneisyltransferase (EC 2.5.1.29) [validated] - Neurospora crassa  
N:Alternate names: geranylgeranyl pyrophosphate synthetase  
C:Species: Neurospora crassa  
C>Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 02-Sep-2000  
C/Accession: S15662; T46596; T46592  
R:Carattoli, A.; Romano, N.; Ballarito, P.; Morelli, G.; Macino, G.  
J. Biol. Chem. 266, 5854-5859, 1991  
A>Title: The Neurospora crassa carotenoid biosynthetic gene (albino 3) reveals highly  
A:Reference number: S15662; M0ID:91170267  
A:Accession: S15662  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-428 <CAR>  
A:Cross-references: GB:U02940; EMBL:X53979; NID:9903318; PIDN:AAC13867.1; PID:9903319  
R:Vittorio, P.; Carattoli, A.; Londel, P.; Macino, G.  
J. Biol. Chem. 269, 26650-26654, 1994  
A>Title: Internal translational initiation in the mRNA from the Neurospora crassa alb  
A:Reference number: A55065; M0ID:95014519  
A:Accession: T46598  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: MEHT/1-129 <VT1>  
A:Cross-references: EMBL:S74011; PIDN:CA833185.1  
A:Accession: T46592  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: MEHT/1-77; PACSOLSRHQMPPSRSLPRTQTLRPK; TSSAPKAPGKRRRR' <VT2>  
A:Cross-references: EMBL:S74685; PIDN:CA833252.1  
A:Experimental source: mutant rosy  
A>Note: translation starts at an alternate initiator and produces a partially active  
C:Genetics:  
A:Gene: al-3  
C:Keywords: transferase

Query Match 50.9%; Score 797.5; DB 2; Length 428;  
Best Local Similarity 50.7%; Pred. No. 2,8e-52;  
Matches 151; Conservative 57; Mismatches 85; Indels 5; Gaps 2;  
DB 6 ETVOIILEPYKYLQIPGKOVRTLSQAFNMLKYPEDKIQIIIEVTEMLHNSLLIDD 65  
123 EEKEKYLTGPDYDLNGHPGKDIRSQWKAFDAMLDVSESLVITTKVISMHTASLLVDD 182  
DB 66 IENSKRKRGEPVAHSTYIGPSVINSANYFFLGLEVLFLDHPDAVKLTROLLHOG 125  
183 VEDNSVLRGEPVAHSTYIGPSVINSANYFFLGLEVLFLDHPDAVKLTROLLHOG 242  
DB 126 OGDLIWRDNYTCTPTEERYKAVYLQKGTGLGVLAVGIMQFSDYKEDLKLTLTLGLFQ 185  
243 QGMDLFWROTLCTPTEDDYLDEMVSNTKGTGLFGLGIMQFSDYKEDLKLTLTLGLFQ 302



[illegible]

RESULT 5  
C71295  
Probable octaprenyl-diphosphate synthase - syphilis spirochete  
C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)  
C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 17-Mar-2000  
C:Accession: C71295  
C:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; McDe  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.  
A:Reference number: A71250; MUID:98332770  
A:Accession: C71295  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-364 <COL>  
A:Cross-references: GB:AE001242; GB:AE000520; NID:33322976; PIDN:AMC65646.1; PID:93322998  
A:Experimental source: strain Nichols  
C:Genetics:  
A:Gene: TP0683  
A:Superfamily: prenyl transferase A

[illegible]

RESULT 6  
C83075  
octaprenyl-diphosphate synthase PA4569 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
C:Species: *Pseudomonas aeruginosa*  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C/Accession: C83075  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbbyg, K.; Lim  
, Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000

A::Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A::Reference number: A82950; MUID:20437337  
A::Accession: C83075  
A::Status: preliminary  
A::Molecule type: DNA  
A::Residues: 1-322 <*SN*>  
A::Cross-references: GB:AE004870; GB:AE004091; NID:99950807; PIDN:MAG07957.1; GSPDB:GN000006  
A::Experimental source: strain PA01  
C::Genetics:  
A::Gene: ispB; PA4569  
C::Superfamily: prenyl transferase A

	Query Match	15.28;	Score 238.5;	DB 2;	Length 322;	
	Best Local Similarity	28.48;	Pred. No. 1.4e-10;			
	Matches	80;	Conservative	53;	Mismatches 130;	Indels 17; Gaps 8;
QY	18 YILLOPGKOVSTTKYSQAENHWLKVPEDEKRLQIIIEVTEMLAHNASLLIDIEDNSKLRRGFP	77				
DQ	38 YIISGGKRRLRPVLVLAGKTLYGKKDDCLLAATIEPLHTSTLIHDDVDVASGLRGRS	97				
QY	78 VAHSIYG-IPSVINSAANYVELGEVKVLTLDHPDAVKLFTRQLLELHGOGGLDI-YWRND	135				
DQ	98 TANLMGWAPSVL-VGDFLYARSEEMWELSGSPVMYRIISATRVINAEGLVLSKYRDA	156				
QY	136 YTCPTEERYKAMVLQKTGTFGL-----AVGMQLPDSYKEDLKPLTLNLSGLEFQIRDDYA	191				
DQ	157 ST--TEELYMEVINGKTMLEFEASTHSAALCOAGEQSALRFREGVLTGAIFQVLDLL	214				
QY	192 NLHRSKEYEKNSSPFCEDLLEGKFSPFTIAHMSRESTVOYNIL-----RQRLENIDIKKYC	247				
DQ	215 DYRGDPAATLVGNVDDILAEGK--PTLPFLVTMRDGEEQALVRKAIQOOGSDLESVC	271				
QY	248 VHYLEDVGSPFYETRTNLKELEAKAYKOIDANGNPDELVALTK	289				
DQ	272 A-AVEAGACADDTANLARDYARAIACDLTLPDNEYRSALVE	312				

```

RESULT      7
T45152
b:functional short chain isoprenyl diphosphate synthase [imported] - Methanobacterium
C:Species: Methanobacterium thermoautotrophicum
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000
C:Accession: T45152
R:Chen, A.; Poulier, C.D.
A:Arch. Biochem. Biophys. 314, 399-404, 1994
A:Title: Isolation and characterisation of ldsA: the gene for the short chain isopren
A:Reference number: 222932; MUID:95070145
A:Accession: T45152
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-325 <CHE>
A:Cross-references: EMBL:S75695; PIDD:AA032421.1
C:Genetics:
A:Gene: ldsA
C:Superfamily: prenyl transferase A

```

	Query Match	14.9%	Score 233.5;	DB 2;	Length 325;
	Best Local Similarity	27.5%;	Pred. No. 3.4e-10;		
	Matches	84;	Conservative	55;	Mismatches 125; Indels 41; Gaps 13;
OY	1	M E K T O E V I R L I L P E Y K Y L L D P G K O V R T K I S - - - Q A F N M L K V P E D K L O I I I E V T M L	56		
		I I : : : : : I I : : : I I : : : I I : : :			
D b	20	M E C I S D I P P T L W A S E H L I T A G C K I R P S L A L L S C E A V G - - - G N P E D A G A C V A A I - E L I	75		
OY	57	H N A S L I D D I E D N S K L R R G P V A N S I Y G I B S V I N S A N T V Y F L G L E K V L T L D H - - - P D A V	112		
		I I : : : I I : : : I I : : : I I : : : I I : : : I I : : :			
D b	76	H T F S L I H D I M D D E M R G E R S P V H I M G E P M A I I A G V L F S K A F E A V I R N D S E R V A L	135		
OY	113	K L F R O L L E L H G O G D I Y W R D N T C P T E E Y K R M V L O K T G L G L A V - - - G L M Q L R S D Y	169		
		I I : : : I I : : : I I : : : I I : : : I I : : :			
D b	136	A V V D S C K I C E G A L D M G F E R L D V - T E D E Y M E M I Y K K T A L L A A T T K A C A I N G A S E R	194		

*Thermoplasma acidophilum* Delta H: funct

```

:| | | : : | : : | : : : : : : || |

```

Db 80 AKTRRGESANLVGNGAVLTGDMYAKSLHFTSTYGNMEMIRLVSGVMDAEGQVLE 139  
 QY 130 IYMRDNTCTEETEEKYKAMVLOKGTGLFGLAVGLMQL---FSYKEDKPLTLNLGLFQOI 166  
 Db 140 ISKVGDI--SEEEYFOIDKGTGLFGLACFVGAMSGEGDMREIYEAQLR-LGRFOL 196  
 QY 187 RDDYANLHSEYSEKSFCEDLTEGKFSPTIHAIMSPRESTQVONILRORTENIDIKKY 246  
 Db 197 IDDALDVEGPKVGLKGVNDLRGKCTYPLI-SVLNDNDREYKRVLRLEDESEKLRK 255  
 QY 247 CVHLEDVGSFEYTRNLKLELAKAKYKQI 275  
 Db 256 VV---ELGGVEKTKERAKE-ELKVKAEI 279

RESULT 11  
 A40433  
 probable pyrophosphate dehydrogenase (crtf) homolog - Cyanophora paradoxa  
 C:Species: Cyanophora paradoxa  
 C:Date: 28-Feb-1992 #sequence\_revision 28-Feb-1992 #text\_change 24-Sep-1999  
 R:Michalowski, C.B.; Loeffelhardt, W.; Bohmert, H.J.  
 J. Biol. Chem. 266, 11866-11870, 1991  
 A:Title: An ORF323 with homology to crtF, specifying prephosphatase pyrophosphate dehydrog  
 A:Reference number: A40433; MUID:91268060  
 A:Accession: A40433  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <MIC>  
 A:Cross-references: GB:M37111; GB:M61174; NID:9336638; PIDN:AAA65472.1; PID:9336639  
 C:Superfamily: prenyl transferase A

Query Match 13.9%; Score 217; DB 2; Length 323;  
 Best Local Similarity 29.8%; Pred. No. 5.8e-09;  
 Matches 67; Conservative 36; Mismatches 108; Indels 14; Gaps 5;

QY 49 IIEVTEMLNASHLLIDIDENSKLRGFPVAHSITGIPSYNSANYVFLGKLVTLIDH 108  
 Db 73 LAETTEIHTASLVHDDIDESDVRKRPVHSDFGKIALIAGDFLFAOSSWYLANLES 132  
 QY 109 PDAVKLETRQLLEHOGQ---GLDIYWRDNYTCPTTEEEKYKAMVLOKGTGLFGLAVGLMQL 165  
 Db 133 LEVYKLSKIVTDPABEIRGLNPFKVD---LTLEVELEKSPYKTSASLAASSKAAAL 188  
 QY 166 FS---DYKEDKPLTLNLGLFQIIRDDYANLHSEYSEKSFCEDLTEGKFSPTIHA 221  
 Db 189 LSHVDLTAVNDLVNNGRHGLAFQIYDDIDFTSSTELGKPCSDLKKGNTLAPVLFAL 248  
 QY 222 WSRPESTQVONILRORTENIDIKKYCVHLEDVGSFEYTRNTLKE 266  
 Db 249 --BQNSLIPLIQROFSEPKDF-EYTLQIVEETKAIETKRELAME 290

RESULT 12  
 T06874  
 probable prenyl transferase (EC 2.5.1.-) - Cyanophora paradoxa cyanelle  
 C:Species: Cyanelle Cyanophora paradoxa  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 24-Sep-1999  
 C:Accession: T06874  
 R:Stewart, V.L.; Michalowski, C.B.; Luffelhardt, W.; Bohmert, H.J.; Bryant, D.A.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: Nucleotide sequence of the cyanelle genome from Cyanophora paradoxa.  
 A:Reference number: 215840  
 A:Accession: T06874  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-323 <STI>  
 A:Cross-references: EMBL:U0821; NID:91016083; PIDN:AAA81217.1; PID:91016130  
 A:Experimental source: strain Pringsheim LB555  
 C:Genetics:  
 A:gene: preA  
 A:genome: cyanelle

C:Superfamily: prenyl transferase A  
 C:Keywords: cyanelle; transferase

Query Match 13.9%; Score 217; DB 2; Length 323;  
 Best Local Similarity 29.8%; Pred. No. 5.8e-09;  
 Matches 67; Conservative 36; Mismatches 108; Indels 14; Gaps 5;

QY 49 IIEVTEMLNASHLLIDIDENSKLRGFPVAHSITGIPSYNSANYVFLGKLVTLIDH 108  
 Db 73 LAETTEIHTASLVHDDIDESDVRKRPVHSDFGKIALIAGDFLFAOSSWYLANLES 132  
 QY 109 PDAVKLETRQLLEHOGQ---GLDIYWRDNYTCPTTEEEKYKAMVLOKGTGLFGLAVGLMQL 165  
 Db 133 LEVYKLSKIVTDPABEIRGLNPFKVD---LTLEVELEKSPYKTSASLAASSKAAAL 188  
 QY 166 FS---DYKEDKPLTLNLGLFQIIRDDYANLHSEYSEKSFCEDLTEGKFSPTIHA 221  
 Db 189 LSHVDLTAVNDLVNNGRHGLAFQIYDDIDFTSSTELGKPCSDLKKGNTLAPVLFAL 248  
 QY 222 WSRPESTQVONILRORTENIDIKKYCVHLEDVGSFEYTRNTLKE 266  
 Db 249 --BQNSLIPLIQROFSEPKDF-EYTLQIVEETKAIETKRELAME 290

RESULT 13  
 E71101  
 probable geranylgeranyl pyrophosphate synthetase - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: E71101  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se  
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: E71101  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-342 <KAM>  
 A:Cross-references: GB:AP000004; NID:93236131; PIDN:BA30171.1; PID:93257488  
 A:Experimental source: strain OT3  
 A:Note: this accession replaces an interim accession for a sequence replaced by Genba  
 C:Genetics:  
 A:gene: PH1072  
 C:Superfamily: prenyl transferase A

Query Match 13.6%; Score 213; DB 2; Length 342;  
 Best Local Similarity 26.7%; Pred. No. 1.2e-08;  
 Matches 88; Conservative 46; Mismatches 134; Indels 62; Gaps 11;

QY 2 EKTQETVQRIILEPYKYLQLPKQVTRKLSQAFNMHKLKVPEDKLOIIEVT--EMILNA 59  
 Db 24 ELIPEKDPRLVLEAARHYPLAGCKRVRPFV--VLTSTEAVGCGPLRAIYPAVAIEILHNY 81  
 QY 60 SLIIDIENSKLRGFPVAHSITGIPSYNSANYVFLGKLVTLIDHDPDAKLFTRQL 119  
 Db 82 SLVHDDIMDETRRGFPVYHRIWGMATLADGLSKAFEAARAEPPEKARVLEY 141  
 QY 120 L-----ELHOGGIDYWRDNYTCPTTEEEKYKAMVLOKGTGLF-----GLAVGLMQLFSD 168  
 Db 142 IVASNMLCEGQARLDEFKKSIV-TIEEYMEWISGCTALFSAKAVGCIIG-----TD 195  
 QY 169 YKEDKPLNL---TLGIFQIRDDYANLHSEYSEKSFCEDLTEGKFSPTIHAISRP 225  
 Db 196 NEEYIKALSMWGRNVGIAFQIMDVLDLIDAEKRLGKPGVSGDIRKGTGLIVAHFENAD 255  
 QY 226 ESQO-----VONILRQRENID-IKKYCVHLEDVGSFEYTRNTL 264  
 Db 256 EKQKORFLKTFGYAGDVKGRGIIIEDIKSDVMEADIDLKRY-----GSIDYAAETA 307  
 QY 265 KELEAKAYKOIDARGNPDELVALVKHLSKM 294

Db 308 KDMIKKA-----NEALRIIPKSKARN 328

## RESULT 14

E85982

octaprenyl diphosphate synthase [imported] - Escherichia coli (strain O157:H7)

C:Species: Escherichia coli

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001

C:Accession: E85982

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamoudis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: E85982

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-323 <STO>

A:Cross-References: GB:AE005174; NID:912517802; PIDN:AG58321.1; GSPDB:GN00145; UMGF:245

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ispb

C:Superfamily: prenyl transferase A

Query Match 13.4%; Score 210.5; DB 2; Length 323;  
Best Local Similarity 23.7%; Pred. No. 1.8e-08;  
Matches 70; Conservative 62; Mismatches 152; Indels 11; Gaps 5;  
QY 1 MEKTOEYVORILLEYKYLLOPGKQVTRKLSQAFNMHLKVPEDKLIITEVTEMLHNAS 60  
Db 22 LEQNSDVOLNOLGY-YIVSGGKRIRPMIAVLAARVGEYGNANTYIALIEFHITAT 80  
QY 61 LLIDIEDNSKLRGFPVAHSITGIPSVNSANYVFLGLEKVTLLDHPDAVKLFTQQL 120  
Db 81 LLHDVDESMDRGKATANAFAFGNASVLDGDTYTRAFQMTSLGSLKYLEVMSAVN 140  
QY 121 ELHOGGGLDIYWRDNYTCP--TEEEYKAMVLOKTGFLGVLAMOLF--DYKEDLK 174  
Db 141 VIAEGEVLQLM--NVNDPDTTEENYMRVIYSKTAFLFEAAQSGSLACTPEEKGLO 197  
QY 175 PLNTLGLFQIRDDYANLHSEKSEKSCEDLTEGKFSPTTHAI-WRPESTOVQNT 233  
Db 198 DYGRYLGTAQRLIDDLIDYNADGOLKNGVDLNEGKPTPLPLHAMHCTPQAOQWIRT 257  
QY 234 LRQTEINIDIKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNPETVALV 288  
Db 258 AIEQNGRHLKPYLEAMNACGSIEMTFRORAEADKAIATALQVLPDTPWREALI 312

## RESULT 15

A81213

octaprenyl diphosphate synthase (EC 2.5.1.-) NMA2161 [imported] - Neisseria meningitidis

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 02-Feb-2001

C:Accession: A81213; F81788

R:Teitelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Science 287, 1809-1815, 2000

A:Authors: Grandl, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A:Reference number: A81000; MUID:20175755

A:Accession: A81213

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-324 <TE7>

A:Cross-References: GB:AE002389; GB:AE002098; NID:97225537; PIDN:AAFA0771.1; PID:9722554

A:Experimental source: serogroup B, strain MC58

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491

A:Accession: F81788

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-324 <PAR>

A:Cross-References: GB:AL162758; GB:AL157959; NID:97380672; PIDN:CA85373.1; PID:9738

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMB0326; NMA2161

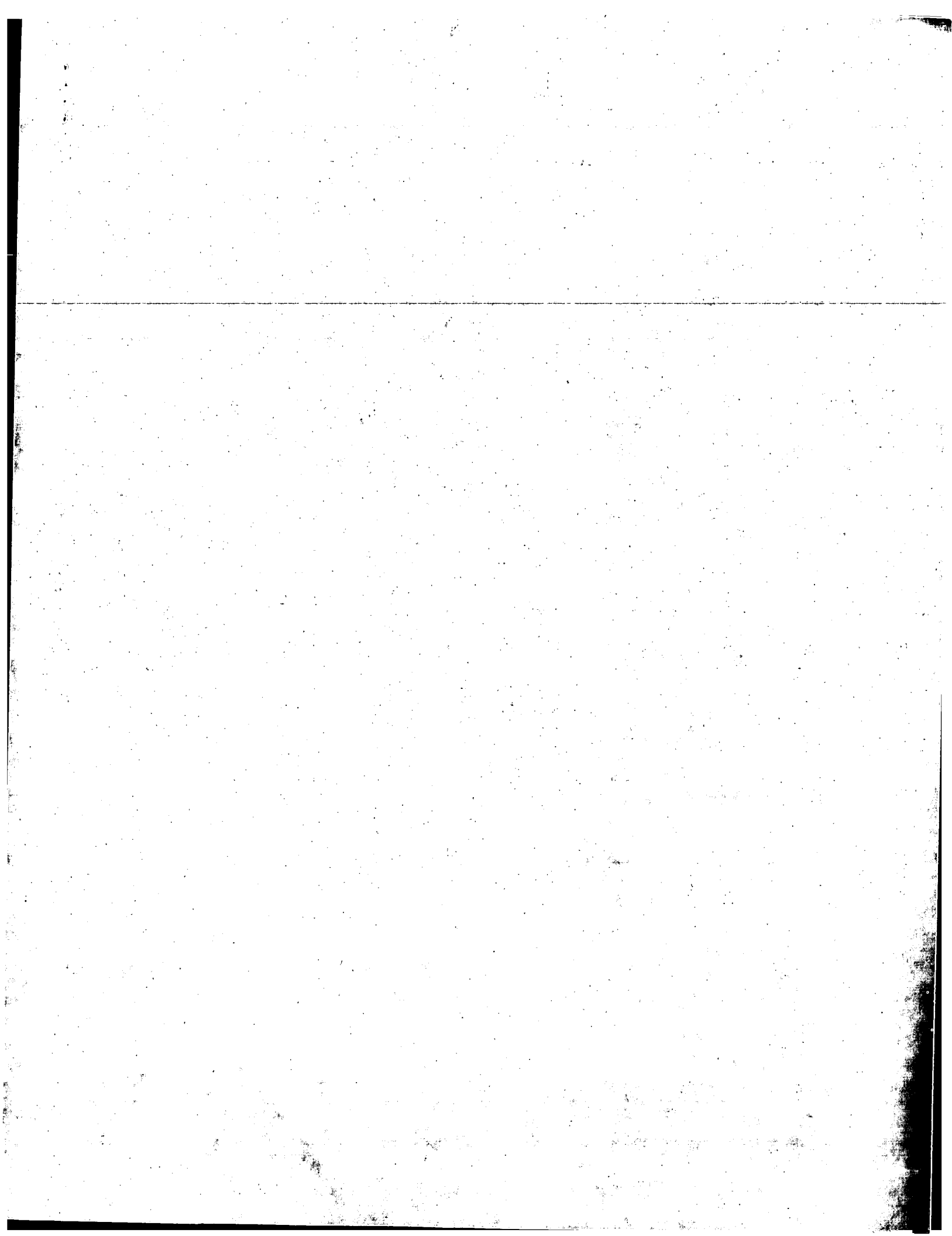
C:Superfamily: prenyl transferase A

C:Keywords: transferase

Query Match 13.4%; Score 209.5; DB 2; Length 324;  
Best Local Similarity 24.4%; Pred. No. 2.1e-08;  
Matches 70; Conservative 61; Mismatches 137; Indels 19; Gaps 6;  
QY 1 MEKTOEYVOR-----LLEPYKYLLOPGKQVTRKLSQAFNMHLKVPEDKLIITEVTE 53  
Db 16 LARVNEVINRAVOSDVALLISQISTYIISAGKRLRPMITILAGKANGVYDEKLYSLAANV 75  
QY 54 EMLNASTLIDIEDNSKLRGFPVAHSITGIPSVNSANYVFLGLEKVTLLDHPDAVK 113  
Db 76 EFTHITSTLLHDVDESMDRGKATANAFAFGNASVLDGDTYTRAFQMTSLGSLKYLEV 135  
QY 114 LFTROLLEHOGGGLDIYWRDNYTCPTEEEYKAMVLOKTGFLGVLAVG--LMOLFSDY 169  
Db 136 VMADATNTIIAEGVWGLMIGN-TDITEBOYIOYQYKAKLFEAAOYGAILGASPEH 194  
QY 170 KEELKPLNTLGLFQIRDDYANLHSEKSEKSCEDLTEGKFSPTTHAIWSRPESTQ 229  
Db 195 ERALKQYGMVGTAFQRLIDDLIDYNADGOLKNGVDLNEGKPTPLPLIYLM--RQSGEQ 252  
QY 230 YONILRQTEINIDIKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQID 276  
Db 253 VANDVTALENNDRS-----YFKIHDIYVRSDALAYISGEARRAVD 294

Search completed: August 19, 2001, 11:21:37  
Job time: 61 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2001, 11:20:36 ; Search time 22.61 seconds  
(Without alignments)  
1755.485 Million cell updates/sec

Title: US-09-744-527-4  
Perfect score: 1566  
Sequence: 1 MEKQETVQRIILEPYKYL.....NPELVATVKhLSKMKKEENE 300

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues  
Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :  
1: SPTEMBL\_16:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_unclassified:\*  
13: sp\_vertebrate:\*  
14: sp\_virus:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	928	59.3	338	5	061539 drosophila
2	928	59.3	338	5	09V554 drosophila
3	862.5	55.1	335	5	061538 drosophila
4	806.5	51.5	303	3	09P885 mucor clrc1
5	709	45.3	392	3	09P965 nigrospora
6	583	37.2	335	3	012051 saccharomyc
7	477.5	30.5	342	3	074289 gibberella
8	310	19.8	347	2	051294 borrellia bu
9	257	16.4	364	2	005708 treponema p
10	254.5	16.3	742	5	09NEB3 leishmania
11	238.5	15.2	322	2	09HVL5 pseudomonas
12	231.5	14.8	299	2	09XIM1 thermotoga
13	231	14.4	304	2	067044 aquifex aeo
14	225.5	14.4	322	2	09R828 pseudomonas
15	213	13.6	342	1	058799 pyrococcus
16	209.5	13.4	324	2	09UOV7 neisseria m
17	201.5	12.9	346	1	09V077 pyrococcus
18	201	12.8	325	2	024743 rhodobacter
19	200	12.8	337	2	09ZD65 rickettsia

20	199	12.7	297	2	09PHX3 campylobact
21	198	12.6	324	2	09S5E9 synecococc
22	197	12.6	332	2	09PD18 xylella fas
23	195.5	12.5	347	1	09H013 halobacteri
24	192.5	12.3	300	1	09HND0 halobacteri
25	192.5	12.3	317	1	09V305 archaeoglob
26	192.5	12.3	326	1	027998 archaeoglob
27	189.5	12.1	300	10	09SYNO archaeobidops
28	189	12.1	272	2	09W08 thermotoga
29	184.5	11.8	288	2	09A302 bacillus fl
30	184	11.7	323	8	09T1S1 cyanidium c
31	182	11.6	348	2	09KUT1 vibrio chol
32	180	11.5	335	2	006428 mycobacteri
33	179.5	11.5	329	2	09RUJ1 micrococcus
34	178.5	11.4	325	2	066129 micrococcus
35	178.5	11.4	379	10	09SHG4 arabidopsis
36	177	11.3	318	1	09UWR6 aeropyrum p
37	177	11.3	337	1	09YB31 aeropyrum p
38	175.5	11.2	291	2	09PFJ6 xylella fas
39	170.5	10.9	344	10	09LJY2 arabidopsis
40	170	10.9	287	2	09FAF1 bacillus su
41	164.5	10.5	430	5	09V923 drosophila
42	163.5	10.4	298	2	09RME5 zymomonas m
43	163	10.4	320	2	09RVU0 deinococcus
44	162.5	10.4	321	10	09FSW8 citrus sine
45	161.5	10.3	422	10	09FT89 arabidopsis

## ALIGNMENTS

RESULT 1  
ID 061539 PRELIMINARY; PRT; 338 AA.  
AC 061539;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE GRANTYGERANYL PYROPHOSPHATE SYNTHASE.  
GN OM OR C68593.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lai C.O., McMahon R., Young C., Mackay T.F.C., Langley C.H.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lai C.O., McMahon R., Young C., Mackay T.F.C., Langley C.H.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF049659; AAC05273.1;  
DR FlyBase; FBgn0019662; qm.  
DR InterPro; IPR000092;  
DR Pfam; PF00348; PolyPrenyl\_synth.1.  
DR PROSITE; PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
SQ SEQUENCE 338 AA; 38860 MW; 0BB7D9F25F2A3D50 CRC64;

Query Match 59.3%; Score 928; DB 5; Length 338;  
Best Local Similarity 59.3%; Pred. No. 2.2e-61;  
Matches 172; Conservative 50; Mismatches 68; Indels 0; Gaps 0;

OY 2 EKQETVQRIILEPYKYLQPGKOVTRKLSQAFNHMLKVPEDKLOITITENTLNASL 61  
Db 14 KSTQKEDDELILQPFYTIQIPKQRFSEIALAFNMHLLIPGEKLAIGIVOMLNHSSL 73  
OY 62 LIDDIEDNSKLRGCPVAHSIYGVINSANYVFLGLEKVTLLDHPDAVKETROLLE 121  
Db 74 LIDDIEDNSILRGVPAVHSIYGVASINANYALFLALEKVOOLDHPDAVKETRYTQLE 133

QY 122 LHOGGDIYWRDNYTCPTTEEEKRANVLOKRTGGLFGLAVGLMQLFSDYKEDLKLPLNTLG 181  
 DB 134 LHRGQGEIYWRDSEFTCPSESDYKLTMTVTKTGLFGLMQLFSSNKEDYKLTALIG 193  
 QY 182 LFFQIRDDYANLHSEKSEKSCEDLTEGKFSFPTTHAHSPESTOVONILRORTENI 241  
 DB 194 LYFOIRDDYCNLSLKEYTEKNSFADLTEGKFGFPVIAHVRTOKOKOVAILHILRQTHDI 253  
 QY 242 DIRKCYVHLEDVGSFPTNTLKELEAKAYKOIDARGNPVELVLYKH 291  
 DB 254 EVKCYCTTLEKLGSGFYTKRVLESIDAEARSEVALGSPNPMIDLNL 303

RESULT 2  
 QYVS54 PRELIMINARY: PRT: 338 AA.  
 AC QYVS54:  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE OM GENE PRODUCT.  
 OS OM OR CG8593.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY.  
 RX MEDLINE-20196006; PubMed-10731132.  
 RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer V., Chumpe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Boltskov S.,  
 RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA de Pablos R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Foster C.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jaisi M., Kaulush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mallet J., McIntosh T.C., Moleod M.P., Mobergson D.,  
 RA Merkulov G., Mishina N.V., Moberg B., Murphy L., Muzny D.M., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclad J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Sotter E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhao X., Zhu S., Zhu X., Smith H.O.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou Q., Zheng L.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 DR EMBL: AF003559; AAF50574.1;  
 DR FlyBase; FBgn0019662; gm.  
 DR InterPro; IPR000092;

DR Pfam: PF00348; polyprenyl\_synth. 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 SQ SEQUENCE 338 AA; 38826 MW; 0BB/7DFCBZ3A3050 CRC64;

Query Match 59.3%; Score 928; DB 5; Length 338;  
 Best Local Similarity 59.3%; Pred. No. 2.2e-61;  
 Matches 172; Conservative 50; Mismatches 68; Indels 0; Gaps 0;

QY 2 EKTQETVORILLPEYKILLDIPGQVTRKLSQAFNMHMKVPEKLOITIEFTEMLNASTL 61  
 DB 14 KSTQKEDDELLQPFYTIQIQPKQFSELAFLFNHMLIPGEKLAQIGDIYQMLNSSL 73  
 QY 62 LIDIEDNSKLRRGPPVAHSITGIPSYNSANVYELGLEKVLTDHPDAVKFTQOLLE 121  
 DB 74 LIDIEDNSKLRRGPPVAHSITGIPSYNSANVYELGLEKVLTDHPDAVKFTQOLLE 133  
 QY 122 LHOGGDIYWRDNYTCPTTEEEKRANVLOKRTGGLFGLAVGLMQLFSDYKEDLKLPLNTLG 181  
 DB 134 LHRGQGEIYWRDSEFTCPSESDYKLTMTVTKTGLFGLMQLFSSNKEDYKLTALIG 193  
 QY 182 LFFQIRDDYANLHSEKSEKSCEDLTEGKFSFPTTHAHSPESTOVONILRORTENI 241  
 DB 194 LYFOIRDDYCNLSLKEYTEKNSFADLTEGKFGFPVIAHVRTOKOKOVAILHILRQTHDI 253  
 QY 242 DIRKCYVHLEDVGSFPTNTLKELEAKAYKOIDARGNPVELVLYKH 291  
 DB 254 EVKCYCTTLEKLGSGFYTKRVLESIDAEARSEVALGSPNPMIDLNL 303

RESULT 3  
 ID 061538 PRELIMINARY: PRT: 335 AA.  
 AC 061538;  
 DT 01-AUG-1998 (TREMBlrel. 07, Created)  
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE GERANYGERANYL PYROPHOSPHATE SYNTHASE.  
 GN OM OR QYEMAO OR CG8593.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SAMARKAND.  
 RX MEDLINE-98085976; PubMed=9426007;  
 RA Lai C., Langley C.H.;  
 RT "A homologue of the 19 kDa signal recognition particle protein locus  
 in Drosophila melanogaster."  
 RL Gene 203:59-63 (1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SAMARKAND.  
 RA Lai C.Q., Langley C.H.;  
 RL Genetics 0:0-0 (1998).  
 DR EMBL: AF049658; AAC05595.1;  
 DR FlyBase; FBgn0019662; gm.  
 DR InterPro; IPR000092;  
 DR Pfam; PF00348; polyprenyl\_synth. 1.  
 DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN 1.  
 SQ SEQUENCE 335 AA; 38391 MW; 72DDID9AF08BFF9030 CRC64;

Query Match 55.18%; Score 862.5; DB 5; Length 335;  
 Best Local Similarity 56.6%; Pred. No. 1.6e-56;  
 Matches 164; Conservative 47; Mismatches 76; Indels 3; Gaps 1;

QY 2 EKTQETVORILLPEYKILLDIPGQVTRKLSQAFNMHMKVPEKLOITIEFTEMLNASTL 61  
 DB 14 KSTQKEDDELLQPFYTIQIQPKQFSELAFLFNHMLIPGEKLAQIGDIYQMLNSSL 73

OY 62 LIDDIENSKLRGCPVAHSHSYGIPSVINSANYVFLGLEKVLTLDPDAVKLTLEHROLLE 121  
 DB 74 L-----TENSLIRKRGVPAHSHITGVASTINANYALFALALEKVOQLDPEKRYTLEOLLE 130  
 OY 122 LHOGGGLDIYWRDNYTCPTTEEEKKAVYLQKTGGLFGLAVGLMOLFSDYKEDLKLPLTLG 181  
 DB 131 LHRGQGEIYWRBSFTCPSSSDYKRLMTVRKTGGLFMLATRLMOLFSSNKEDYSKLTALIG 190  
 OY 182 LFPQIRDDYANLHSHKESEKSCFEDLTEGKFSPTTHAHSRPESTOVONILRQRTENT 241  
 DB 191 LFPQIRDDYCNLSLKEYSVKSFEADLTEGKFGFPYIHAVTKODKQVLRKSSQRTHDI 250  
 OY 242 DIKKYCVHYLEDVGSFEYTRNTLKELEKAKAYKQDARGSDPELVALVKHL 291  
 DB 251 EVKKYCTITLLEKLGSSFOYTRKRVLESIDAEARSEYSTLGSPMYDRLLNKL 300  
 RESULT 4  
 O9P85 PRELIMINARY; PRT; 303 AA.  
 AC 09P85;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHASE (EC 2.5.1.29).  
 GN CARG.  
 OS Mucor circinelloides f. lusitanicus.  
 OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;  
 OC Mucor.  
 OX NCBI\_TaxID=29924;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS277.49;  
 RA Velazos A., Fuentes M., Aguilar R., Eslava A.P., Iturriaga E.A.;  
 RT "Prenyl synthases in Mucor circinelloides.";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ276129; CAB89115.1;  
 DR Interpro: IPR000092;  
 DR Pfam: PF00348; Polyprenyl\_synth; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; UNKNOWN\_1.  
 DR POSTSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN\_1.  
 KM transferase.  
 SQ SEQUENCE 303 AA; 34787 MW; D440757CA7D0E215 CRC64;

Query Match 51.5%; Score 806.5; DB 3; Length 303;  
 Best Local Similarity 53.4%; Pred. No. 2e-52;  
 Matches 151; Conservative 52; Mismatches 79; Indels 1; Gaps 1;

OY 3 KTOE-TYQRIILEPKYKLLQIPKQVTKLSQAFNHLKVPEDKQIITIEVTEMLHNASL 61  
 DB 7 RTEERSTEDITILEPYTLISQPGDIRAKLISAFDLWHPKVDLCVINKTIGMLHNASL 66  
 OY 62 LIDDIENSKLRGCPVAHSHSYGIPSVINSANYVFLGLEKVLTLDPDAVKLTLEHROLLE 121  
 DB 67 MIDDVQDSDLRKGVPAHSHITGVASTINANYALFALALEKVOQLDPEKRYTLEOLLE 130  
 OY 122 LHOGGGLDIYWRDNYTCPTTEEEKKAVYLQKTGGLFGLAVGLMOLFSDYKEDLKLPLTLG 181  
 DB 127 LHRGQGEIYWRBSFTCPSSSDYKRLMTVRKTGGLFMLATRLMOLFSSNKEDYSKLTALIG 190  
 OY 182 LFPQIRDDYANLHSHKESEKSCFEDLTEGKFSPTTHAHSRPESTOVONILRQRTENT 241  
 DB 187 LHRGQGEIYWRBSFTCPSSSDYKRLMTVRKTGGLFMLATRLMOLFSSNKEDYSKLTALIG 190  
 OY 242 DIKKYCVHYLEDVGSFEYTRNTLKELEKAKAYKQDARGSDPELVALVKHL 291  
 DB 247 EVKKYCTITLLEKLGSSFOYTRKRVLESIDAEARSEYSTLGSPMYDRLLNKL 300

RESULT 5  
 O9P965

ID O9P965 PRELIMINARY; PRT; 392 AA.  
 AC O9P965;  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)  
 DE GERANYLGERANYL DIPHOSPHATE SYNTHASE.  
 OS Nigrospora sphaerica.  
 OC Eukaryota; Fungi; Ascomycota; mtosporic Ascomycota; Nigrospora.  
 OX NCBI\_TaxID=114231;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBS 984.69;  
 RA Ohashi K., Sankawa U.;  
 RT "Nigrospora sphaerica geranylgeranyl diphosphate synthase.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB037600; BAA90525.1;  
 DR Interpro: IPR000092;  
 DR Pfam: PF00348; Polyprenyl\_synth; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; UNKNOWN\_1.  
 DR POSTSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN\_1.  
 SQ SEQUENCE 392 AA; 43615 MW; DDA1A5196864AA43 CRC64;

Query Match 45.3%; Score 709; DB 3; Length 392;  
 Best Local Similarity 46.7%; Pred. No. 5.1e-45;  
 Matches 156; Conservative 51; Mismatches 100; Indels 4; Gaps 1;

OY 6 EYQRIILEPKYKLLQIPKQVTKLSQAFNHLKVPEDKQIITIEVTEMLHNASL 65  
 DB 99 EAKKVVYTGYPDYIAASPGKIRSLMLACFNAMLEVPQAEIRKRAVGLMHTFASLID 158  
 OY 66 IEDNSKLRGCPVAHSHSYGIPSVINSANYVFLGLEKVLTLDPDAVKLTLEHROLLE 121  
 DB 159 IODNSKLRGCPVAHSHSYGIPSVINSANYVFLGLEKVLTLDPDAVKLTLEHROLLE 121  
 OY 126 LIDDIENSKLRGCPVAHSHSYGIPSVINSANYVFLGLEKVLTLDPDAVKLTLEHROLLE 121  
 DB 219 QAMDLYRRELTLCTEADYEMTSNKTGGLFRLAYRLMKESANSVDLMVVELLGLVLFQ 278  
 OY 186 IRDDYANLHSHKESEKSCFEDLTEGKFSPTTHAHSRPESTOVONILRQRTENT 241  
 DB 279 VADYKMLCSREYDGLDGVEDLTEGKFSPTTHAHSRPESTOVONILRQRTENT 241  
 OY 246 YCVHYLEDVGSFEYTRNTLKELEKAKAYKQDARGSDPELVALVKHL 291  
 DB 339 YALEMKTGTSFRTYTRKRVLESIDAEARSEYSTLGSPMYDRLLNKL 300

RESULT 6  
 O12051 PRELIMINARY; PRT; 335 AA.  
 AC O12051;  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)  
 DE GERANYLGERANYL DIPHOSPHATE SYNTHASE LPELP.  
 GN LPELP OR BTSL.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TRANSPOSON-TY4;  
 RA Winnett E., Ahmed A., Bussey H., Fortin N., Friesen J.D., Hall J.,  
 RA Storms R.K., Vo D.H., Wang Y.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95394944; PubMed=7665600;  
 RA Jiang Y., Proteau P., Poultier D., Ferro-Novick S.;  
 RT "BTSL encodes a geranylgeranyl diphosphate synthase in Saccharomyces cerevisiae.";

```

Query Match      30.5% ; Score 477.5 ; DB 3 ; Length 342 ;
Best Local Similarity 43.2% ; Pred. No. 6,9e-28 ;
Matches 99 ; Conservative 37 ; Mismatches 86 ; Indels 7 ; Gaps 2 ;
Y 15 PYKYLDTLPGKQVPRFKLSQAFNRHMLKVPEDKTLIIIVETMLHNASLLIDIEDNSKLR 74

```

[illegible]

OY 251 LEDVGSFETRNKLEAKAYQIDARGNPELVALVKHLSMPKEE 298  
DB 296 TEMINS-----SKSIR-----NSTIVAL-KYLINE-FKNE 322

RESULT 9  
ID 005708 PRELIMINARY; PRT; 364 AA.  
AC 005708;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)  
DE OCTAPRENYL-DIPHOSPHATE SYNTHASE.  
GN TP0683.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_Taxid=160;  
RA [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NICHOLS.  
RL Steiner B.M., Rodas B., Liu H.;  
RN Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
RP SEQUENCE FROM N.A.  
RC STRAIN=NICHOLS;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Dodergen E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uterback T.,  
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
RT spirochete."  
RL Science 281:375-388(1998).  
DR EMBL: U95214; AAC53642.1;  
DR EMBL: AEO01242; AAC53646.1;  
DR TIGR: TP0683;  
DR InterPro: IPR000092;  
DR Pfam: PF00348; polyrenyl\_syn; 1.  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN\_1.  
SQ SEQUENCE 364 AA; 40170 MW; EB3588FB570D6A69 CRC64;

Query Match 16.4%; Score 257; DB 2; Length 364;  
Best Local Similarity 35.5%; Pred. No. 1.9e-11;  
Matches 70; Conservative 30; Mismatches 85; Indels 12; Gaps 5;

OY 40 KVEDKQIIIEVTEMLHNSLLIDIEDNSKLRGFPVAHSYIGIPSVINSANYVELG 99  
DB 89 KISEREYVTLTSAVELLHTSLHDIEDHSSYRGRKPCAYLRGFCVNAASWLYEHA 148  
OY 100 LEVVLTLHDHDAVK--LEFPOQL--LEHOGGLDIYW-RDNTTCPEEYKAMVLOKTG 133  
DB 149 HALIDTLSTIEPAKALFAFSATISATRALHGOALDIAMHSPELIPSRADYLRVALKTG 208  
OY 154 GLFGLVAGLQLQFSDYK---EDLKPRLNTLGLFQIORDYANLHNSKEYSENSCEDLT 209  
DB 209 ALALSGELGFLCAGYSARARSGAGVMPIGIGFOILDDVQIDG--VSGRAHADDIY 266  
OY 210 EGKFSFPTTHAISRPE 226  
DB 267 EGKFSFPTTHAISRPE 283

RESULT 10  
OYNEB3 PRELIMINARY; PRT; 742 AA.  
AC OYNEB3;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)  
DE POSSIBLE GERANTLGERANTL DIPHOSPHATE SYNTHASE.  
GN L6066.04.  
OS Leishmania major.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_Taxid=5664;  
RA [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PAOI.  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PAOI;  
RX MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
RA Smith D.F.;  
RT "A physical map of the Leishmania major Friedlin genome."  
RL Genome Res. 8:135-145(1998).  
DR EMBL: AL157416; CAB75644.1;  
DR InterPro: IPR000092;  
DR Pfam: PF00348; polyrenyl\_syn; 1.  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
SQ SEQUENCE 742 AA; 82727 MW; F496192F1F51B5B8 CRC64;

Query Match 16.3%; Score 254.5; DB 5; Length 742;  
Best Local Similarity 32.3%; Pred. No. 7.6e-11;  
Matches 95; Conservative 44; Mismatches 118; Indels 37; Gaps 12;

OY 2 EKTQEVYQRLLEPKYKYLDPKQVYTRKLSQAFNHKLPKEDKQIIIEVTEMLHNSL 61  
DB 429 DKICE---LEFRVRSIIDRGKSMRSLIVSCNALSRQYDCRRYIAVELLHVSL 484  
OY 62 LIDIEDNSKLRGFPVAHSYIGIPSVINSANYVELGLEKVTLDHPD---AVKLF 116  
DB 485 LIDIEDNSKLRGFPVAHSYIGIPSVINSANYVELGLEKVTLDHPD---AVKLF 544  
OY 117 ROLLEHOGGLDIYWRDNTTCPEEYKAMVLOKTGGLF---LAVGL 162  
DB 545 DVLLAGAGGGIDIT-RDYLMPKRYETGASLTFDALDAIHRYKCGANGALCMAVCL 603  
OY 163 MQLFSDYKEDKLPNTLGLFQIORDYANLHNSKEYSEN-KSCCEDITEGKFSFPTTHAI 221  
DB 604 CEAPTVLSEAVVERFGLAGLAFQIVDALNI--RGFEENKEEAEDIKDKKITPTAIAM 661  
OY 222 WSPRESQVON---ILRORT-ENIDIKKCVHLEDVGS---FETRNKLEK 267  
DB 662 -GRLEADROTALNLRKPKTEKADIDQ-AVELIMKVDAVSECLARHRLQEM 713

RESULT 11  
OYNEB3 PRELIMINARY; PRT; 322 AA.  
AC OYNEB3;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)  
DE OCTAPRENYL-DIPHOSPHATE SYNTHASE.  
GN ISPB OR PA4569.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OX Pseudomonas.  
RA [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PAOI;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,



QY 247 CVHLEDVGSFEYRTNLTKELEAKAYK01 275  
 DB 256 VV-----ELGVEKTERAKE-ELKVKKEI 279

## RESULT 14

Q9RBZ8 PRELIMINARY; PRT: 322 AA.

AC 09RBZ8; 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)  
 DE OCTYPRENYL DIPHOSPHATE SYNTHASE-LIKE PROTEIN.  
 OS Pseudomonas sp. BG33R.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_Taxid=96901;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BG33R;  
 RA Glandorf D.C.M., Wechter W.P., Derrick W.C., Leverenz B.,  
 RA Kluepfel D.A.;  
 RT "Identification of genetic loci in a rhizosphere inhabiting  
 RT fluorescent Pseudomonas sp. BG33R involved in expression of a  
 RT phytoparasitic nematode ovicidal factor."  
 RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF153713; AAD47627.1;  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; polyprenyl\_synth. 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET 1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET 2; 1.  
 SO SEQUENCE 322 AA; 34904 MW; EA479BD3120B9394 CRC64;

Query Match 14.4%; Score 225.5; DB 2; Length 322;  
 Best Local Similarity 28.1%; Pred. No. 3.5e-09;  
 Matches 80; Conservative 55; Mismatches 127; Indels 23; Gaps 8;

QY 18 YLQLPGKQVRLTSLQAFNHLKVPEDKLOITTEVEMLNHSLDIEDNSKLRGFP 77  
 DB 38 YLTSAGSKRLPLLVLCGLGREGDRLRLATTEFLTATLHDVDVDMGMRGRE 97  
 QY 78 VHSIYV-IPSVINSANYVYFLEKLVLTLDHPDAVKLFTROLLEHOGGLDI-YWRDN 135  
 DB 98 TANAAMGNAPSVL-VGDEMTSRSEFMWELGSMVPVKILISQATRIITAEGLDLSKVRDA 156  
 QY 136 YTCPTTEEEKKAMVLOKTLGFLGFL-ANGLMQLFSDYKEDKPLNLTGLFPOIRDDYA 191  
 DB 157 ST--TEETVMEVIRKTAFFEAHSTSAALCGATAQAALRTFGDNLGVATQVLDL 214  
 QY 192 NLSKSEKSEKSCEDLTDEKFSFPTTHAIWSPESIQVONITRQRTENIDIKKYCHYL 251  
 DB 215 DYKDAETLCKNVDDLADECK--PTLPLIYTRREGTFPEQALVRKA---IQKGIEDL 267  
 QY 252 EDV-----GSFEYTRNTLKELEAKAYKOIDARGGNPDELVALVK 289  
 DB 268 EAIRAVEASGSLFETROLARDYVARAIQCLEALPASEYDPAIVE 312

## RESULT 15

Q58799 PRELIMINARY; PRT: 342 AA.

AC 058799; 01-AUG-1998 (TREMblrel. 07, Created)  
 DT 01-AUG-1998 (TREMblrel. 07, last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)  
 DE 342AA LONG HYPOTHEICAL GERANYLGERANYL PYROPHOSPHATE SYNTHETASE.  
 GN PH1072.  
 OS Pyrococcus horikoshii.  
 OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
 OX NCBI\_Taxid=53953;  
 RN [1]

RP SEQUENCE FROM N.A.

RC MEDLINE-98344137; PubMed-9679194;  
 RA Kawarayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
 RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
 RA Masuchi Y., Shizuya H., Kikuchi H.;  
 RT "Complete sequence and gene organization of the genome of a hyper-  
 RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3."  
 RL DNA Res. 5:55-76(1998).  
 DR EMBL: AF000004; BAA30171.1;  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; polyprenyl\_synth. 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET 1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET 2; 1.  
 SO SEQUENCE 342 AA; 38551 MW; 54DBE110A37704DA CRC64;

Query Match 13.6%; Score 213; DB 1; Length 342;  
 Best Local Similarity 26.7%; Pred. No. 3.2e-08;  
 Matches 88; Conservative 46; Mismatches 134; Indels 62; Gaps 11;

QY 2 EKTQETVORILPEPYKYLQPGKQVRLTSLQAFNHLKVPEDKLOITTEVEMLNHSLDIEDNSKLRGFP 119  
 DB 24 ELIPEKDPRLVLEAARIRYPLAGKRRAPV--VLTSEAVGDPRLAIRYVAVALIELIHNY 81  
 QY 60 SLIIDIEDNSKLRGFPVAHSIYVIGIPSVINSANYVYFLEKLVLTLDHPDAVKLFTROL 119  
 DB 82 SLVHDDIMDDERIRGKPTVHRIMGVMMALADGLLESKAFEAVARAEIPEPKARVLEY 141  
 QY 120 L-----ELHOGGLDIYWRDNYTCPTTEEEKKAMVLOKTLGFL-----GLAVGLMQLFSD 168  
 DB 142 IVKASNELCEGOARDEFEKSTV-TIEEYMEMISGKTALFEASAKVGGIIG--TD 195  
 QY 169 YKEDLKPILN--TIGLFPQIRDDYANLHSEKSEKSCFEDLTDECKFSFPTTHAIWSP 225  
 DB 196 NEEYIKALSSWGRNVGIAFIMDDVLDLIDDEKLGKPVGSDIRKGGKTLIVAHFENAD 255  
 QY 226 ESTQ-----VONILRQRTENID-IRKYCVHLEDVGSFEYTRNTL 264  
 DB 256 EKDKORFLKIFGVYAGDVKRGIIIEDIKSDVMEALDLKKY-----GSDIYAAEIA 307  
 QY 265 KELEAKAYKOIDARGGNPDELVALVKHLSK 294  
 DB 308 KDMIKRA-----NEALRLIPKSKARM 328

Search completed: August 19, 2001, 11:22:08  
 Job time: 92 sec





Result No.	Score	Query Match	Length	DB	ID	Description
1	13	4.3	303	3	Q9P885	Q9P885 mucor circii
2	11	3.7	338	5	O61539	O61539 dirosophila
3	11	3.7	338	5	Q9V654	Q9V654 dirosophila
4	11	3.7	392	3	O9P965	O9P965 nigrospora
5	10	3.3	335	3	O12051	O12051 saccharomyce
6	9	3.0	205	2	O45496	O45496 bacillus su
7	8	2.7	161	2	Q9RKB2	Q9RKB2 streptococ
8	8	2.7	267	2	O9LBG2	O9LBG2 leifsonia a
9	8	2.7	277	4	O9NPC3	O9NPC3 homo sapien
10	8	2.7	335	5	O61538	O61538 dirosophila
11	8	2.7	347	2	O51294	O51294 borrella bu
12	8	2.7	606	5	O18869	O18869 caenorhabdi
13	8	2.7	844	14	O9UH67	O9UH67 sheep astrc
14	8	2.7	1790	3	O07380	O07380 saccharomyce
15	7	2.3	41	14	Q9WS35	Q9WS35 rabies viru
16	7	2.3	62	2	O9XAG1	O9XAG1 streptomyce
17	7	2.3	105	5	O02290	O02290 caenorhabdi
18	7	2.3	112	8	O21620	O21620 davevskia a
19	7	2.3	112	8	O21622	O21622 davevskia d

20	7	2.3	113	5	09NHG1	09bhl1	plasmodium
21	7	2.3	115	8	037739	037739	felis silive
22	7	2.3	138	2	09X2M1	09x2m1	pseudomonas
23	7	2.3	156	10	09SR25	09sr25	arabidopsis
24	7	2.3	174	2	09KFH6	09kfhe	bacillus ha
25	7	2.3	181	5	09NFP8	09nfp8	plasmodium
26	7	2.3	183	2	09WY15	09wy15	thermotoga
27	7	2.3	184	14	09DV95	09dv95	human
28	7	2.3	185	2	09R8R4	09r8r4	helicobacte
29	7	2.3	193	10	09M538	09m538	alonsoa meze
30	7	2.3	198	8	021627	021627	darevska m
31	7	2.3	198	8	09MNY4	09mny4	lacerta arm
32	7	2.3	198	8	09MNT3	09mny3	lacerta dah
33	7	2.3	198	8	09MM68	09mm68	darevska d
34	7	2.3	198	8	09MDL2	09mdl2	darevska m
35	7	2.3	198	8	021629	021629	darevska c
36	7	2.3	204	4	09H5H0	09h5h0	homo sapien
37	7	2.3	218	5	09VQ03	09vq03	trossophila
38	7	2.3	225	3	09HGN3	09hgn3	schizosach
39	7	2.3	241	5	09VDR0	09vdr0	trossophila
40	7	2.3	245	9	09FZT9	09fzt9	pseudomonas
41	7	2.3	247	2	09KT16	09kt16	viديو chol
42	7	2.3	249	2	09RKH4	09rkh4	streptomyce
43	7	2.3	252	1	006104	006104	methanococc
44	7	2.3	258	5	09VS38	09vs38	trossophila
45	7	2.3	273	10	043381	043381	arabidopsis

## ALIGNMENTS

## RESULT 1

ID	PRELIMINARY;	PRT;	303 AA.
AD	Q9P885;		
DT	01-OCT-2000 (TREMblrel_15, Created)		
DT	01-OCR-2000 (TREMblrel_15, Last sequence update)		
DT	01-MAR-2001 (TREMblrel_16, Last annotation update)		
DE	GEANTLGERANYL PYROPHOSPHATE SYNTHASE (EC 2.5.1.29).		
GN	CARG.		
OS	Mucor circinelloides f. lusitanicus.		
OC	Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae		
OC	Mucor		
OX	NCBI_TaxID=29924;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CBS277.49;		
RA	Velasos A., Fuentes M., Aguilar R., Esilava A.P., Iturrigaga E.A.;		
RT	"Prenyl synthases in Mucor circinelloides.";		
RL	Submitted (NPR-2000) to the EMBL/Genbank/DBJ databases.		
DR	EMBL; AJ276129; CAB89115.1; -		
DR	InterPro: IPR000092; -		
DR	Pfam: PFE00348; poly(prenyl)_syn1.1.		
DR	PROSITE: PS00723; POLYPRENYL_SYNTHET_1; UNKNOWN_1.		
DR	PROSITE: PS00444; POLYPRENYL_SYNTHET_2; UNKNOWN_1.		
KW	Transferase.		
QC	SEQUENCE 303 AA; 34787 MW; D440757CA7DDE2J5 CRC64;		

Query Match	4.38;	Score 13;	DB 3;	Length 303;
Best Local Similarity	100.00;	Pred. No. 7.9e-05;		
Matches 13; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
Qy 204 FCEDLTEGKFSFP 216
    |||||
Db 209 FCEDLTEGKFSFP 221
```

RESULT	2		
061539			
ID	061539	PRELIMINARY;	PRT; 338 AA
AC	061539;		
DT	01-AUG-1998	(TREMUR, 07, Created)	

DT 01-AUG-1998 (Tremblrel. 07, last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHASE.  
 GN OM OR CG8593.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Lai C.O., Langley C.H.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RP SEQUENCE FROM N.A.  
 RA Lai C.O., McMahon R., Young C., Mackay T.F.C., Langley C.H.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: A004659; AAC05273.1;  
 DR FlyBase: FBgn0019662; qm.  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; polyprenyl\_synth; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
 SO SEQUENCE 338 AA; 38860 MW; 0BB7D9F25F2A3D50 CRC64.

Query Match 3.7%; Score 11; DB 5; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 0.0096;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 SLIDDIEDNS 70  
 DB 72 SLIDDIEDNS 82

RESULT 3  
 ID 09VS54 PRELIMINARY; PRT; 338 AA.  
 AC 09VS54.  
 DT 01-MAY-2000 (Tremblrel. 13, Created)  
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE OM GENE PRODUCT.  
 GN OM OR CG8593.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Wortman J.R., Lander M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer J.R., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abil J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burdits K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Helman T.J., Wei M.-H., Ibeagwu C.,  
 RA Jalali M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasho P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spletter E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003559; AF050574.1;  
 DR FlyBase: FBgn0019662; qm.  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; polyprenyl\_synth; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 SO SEQUENCE 338 AA; 38826 MW; 0BB7D9FCBF2A3D50 CRC64;

Query Match 3.7%; Score 11; DB 5; Length 338;  
 Best Local Similarity 100.0%; Pred. No. 0.0096;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 SLIDDIEDNS 70  
 DB 72 SLIDDIEDNS 82

RESULT 4  
 ID 09P965 PRELIMINARY; PRT; 392 AA.  
 AC 09P965.  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
 DE GERANYLGERANYL DIPHOSPHATE SYNTHASE.  
 GN Nigrospora sphaerica.  
 OS Nigrospora sphaerica.  
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Nigrospora.  
 OX NCBI\_TaxID=114231;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CBS 984.69;  
 RA Ohashi K., Sankawa U.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB037600; BAA90525.1;  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; polyprenyl\_synth; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
 SO SEQUENCE 392 AA; 43615 MW; D0A1A5196864AA43 CRC64;

Query Match 3.7%; Score 11; DB 3; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 206 EDLTKGFSEFP 216  
 DB 299 EDLTKGFSEFP 309

RESULT 5  
 ID 012051 PRELIMINARY; PRT; 335 AA.

```

AC 012051;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE GENANLYGERANTL DIPHOSPHATE SYNTHASE LPEIP.
GN LPEIC OR BTL1.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSOM-RT4;
RA Winnett E., Ahmed A., Bussey H., Fortin N., Friesen J.D., Hall J.,
RA Storms R.K., Vo D.H., Wang Y.,
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95394944; PubMed=7665600;
RA Jiang Y., Proteau P., Poulter D., Ferro-Novick S.;
RT "BTL1 encodes a geranylgeranyl diphosphate synthase in Saccharomyces
RT cerevisiae."
RL J. Biol. Chem. 270:21793-21799(1995).
DR EMBL; U39205; AAB68296.1;
DR EMBL; U31632; AAA83262.1;
DR SGD: S0005990; BTL1.
DR InterPro: IPR000092;
DR Pfam; PF00348; PolyPrenyl_synth; 1.
DR PROSITE; PS00723; POLYPRENYL_SYNTHET_1; 1.
DR PROSITE; PS00443; POLYPRENYL_SYNTHET_2; UNKNOWN_1.
SQ SEQUENCE 335 AA; 38651 MW; 4C7D6527FE29F157 CRC64;
Oy 60 SLIIDDIEDN 69
Db 71 SLIIDDIEDN 80

Query Match 3.3%; Score 10; DB 3; Length 335;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 SLIIDDIEDN 69
Db 71 SLIIDDIEDN 80

RESULT 6
O45496 PRELIMINARY: PRT: 205 AA.
AC O45496;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)
DE HYPOTHETICAL 23.9 KDA PROTEIN.
GN YKRC OR YKTA.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-1168;
RX MEDLINE=90368558; PubMed=1697575;
RA Hemla H., Palva A., Paulin L., Arvidson S., Palva I.;
RT "Secretory S complex of Bacillus subtilis: sequence analysis and
RT identity to pyruvate dehydrogenase."
RL J. Bacteriol. 172:5052-5063(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-1168;
RX MEDLINE=97124187; PubMed=8969500;
RA Winters P., Caldwell R., Enfield L., Ferrari E.;
RT "The ampr-nprE (124 degrees-127 degrees) region of the Bacillus
RT subtilis 168 chromosome: sequencing of a 27 kb segment and
RT identification of several genes in the area."
RL Microbiology 142:0-0(0).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-1168;
RX MEDLINE=97144523; PubMed=8990290;
RA Henriques A.O., Bryan E.M., Beall B.W., Moran C.P. Jr.;
RT "cseI, cse6, and csk22 are new members of mother-cell-specific
RT sporulation regulons in Bacillus subtilis."
RL J. Bacteriol. 179:389-398(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolyotin A., Borchert S.,
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denliot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Hatech J., Hatwood C.R., Henaut A.,
RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koelter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Manuel C., Medigue C.,
RA Medina N., Meliello R.P., Mizuno M., Moesti D., Nakai S., Nodack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogilwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Portwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
DR EMBL; AF012285; AAC24931.1;
DR EMBL; Z99111; CAB13330.1;
KW Hypothetical protein.
FT CONFLICT 77
SQ SEQUENCE 205 AA; 23912 MW; 5C9A604C240E91C8 CRC64;

Query Match 3.0%; Score 9; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 0.67;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 262 NTKKLEAK 270
Db 36 NTKKLEAK 44

RESULT 7
O9RKB2 PRELIMINARY: PRT: 161 AA.
AC O9RKB2;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
DE PUTATIVE ACETYLTRANSFERASE.
GN SCE87.26.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE-97000351; PubMed-8843436;  
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL132674; CAB59671.1; -;  
 DR InterPro; IPR000182; -;  
 DR Pfam; PF00583; Acetyltransf. 1.  
 KW Transferase.  
 SQ SEQUENCE 161 AA; 17961 MW; 6C998C684B45DC48 CRC64;

Query Match 2.7%; Score 8; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 5.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 107 DHPDAVKL 114  
 DB 9 DHPDAVKL 16

RESULT 8  
 ID O9L8G2 PRELIMINARY; PRT; 267 AA.  
 AC O9L8G2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE LEVODIONE REDUCTASE.  
 GN LVR.  
 OS Leifsonia aquatica.  
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 CC Actinomycetales; Micrococciaceae; Microbacteriaceae; Leifsonia.  
 OX NCBI\_TaxID=144185;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-M-13;  
 RA Wada M., Yoshizumi A.;  
 RL "Levodione reductase gene from corynebacterium aquaticum M-13.";  
 CC Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)  
 CC FAMILY.  
 DR EMBL; AB042262; BA95121.1; -;  
 DR InterPro; IPR002198; -;  
 DR InterPro; IPR002347; -;  
 DR Pfam; PF00106; adh\_short; 1.  
 DR Pfam; PF00678; adh\_short\_C2; 1.  
 DR PRINTS; PR00081; GDHRDH.  
 DR PRINTS; PR00080; SDRFAMILY.  
 KW Oxidoreductase.  
 SQ SEQUENCE 267 AA; 27920 MW; 6F05C89383500304 CRC64;

Query Match 2.7%; Score 8; DB 2; Length 267;  
 Best Local Similarity 100.0%; Pred. No. 8.9;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 97 FLGLEKVL 104  
 DB 129 FLGLEKVL 136

RESULT 9  
 ID Q9NPC3 PRELIMINARY; PRT; 277 AA.  
 AC Q9NPC3;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOTHEICAL 31.5 KDA PROTEIN (ENHANCER OF INVASION 10).  
 GN DKE2P564A0772 OR HE110.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Blum H., Bauersachs S., Mewes H.W., Weil B., Wiemann S.;  
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Golemis E.A., Estojak J., Toby G.G.;  
 RT "The HE110 (human enhance of filamentation) gene induces yeast cell  
 RT polarization.";  
 RL Submitted (DEC-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-HUMAN SMALL INTESTINE;  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Obaayashi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isoigai T., Sugano S.;  
 RT "NEDD human cDNA sequencing project.";  
 RL Submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL161994; CAB82326.1; -;  
 DR EMBL; AF216381; AAF6386.1; -;  
 DR EMBL; AK026233; BAB15403.1; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 277 AA; 31544 MW; 885723B01A35225A CRC64;

Query Match 2.7%; Score 8; DB 4; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 9.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 EBYKAVYL 149  
 DB 70 EBYKAVYL 77

RESULT 10  
 ID O61538 PRELIMINARY; PRT; 335 AA.  
 AC O61538;  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE GRANULIGERANYL PYROPHOSPHATE SYNTHASE.  
 GN OM OR QJEMO OR CG8593.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SAMARKAND;  
 RA MEDLINE-98085976; PubMed-9426007;  
 RA Lal C., Langley C.H.;  
 RT "A homologue of the 19 kDa signal recognition particle protein locus  
 RL in Drosophila melanogaster.";  
 RN Gene 203:59-63(1997).  
 RP SEQUENCE FROM N.A.

RC STRAIN-SAMARKAND;  
 RA Lal C.O., Langley C.H.;  
 RL Genetics 0:0-0(1998).  
 DR EMBL: AF049658; AAC05595.1;  
 DR FLYBase: FBgn0019662; gnm.  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
 SQ SEQUENCE 335 AA; 38391 MW; 72D1D9AF08EF9030 CRC64;

Query Match 2.7%; Score 8; DB 5; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 77 PYAHSIY 84  
 Db 86 PYAHSIY 93

RESULT 11  
 O51294 PRELIMINARY; PRT; 347 AA.  
 AC O51294;

DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)  
 DE OCTAPRENYL-DIPHOSPHATE SYNTHASE (ISBP).

GN BB0314.  
 OS Borrelia burgdorferi (Lyme disease spirochete).  
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
 OX NCBI\_TaxID=139;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 35210 / B31;  
 RX MEDLINE=96065943; PubMed=9403685;

RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
 RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
 RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
 RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
 RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
 RA Uitterlind T., White L., McDonald L., Artiach P., Bowman C.,  
 RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
 RA Smith H.O., Venter J.C.;

RT "Genomic sequence of a Lyme disease spirochete, Borrelia  
 burgdorferi."  
 RL Nature 390:580-586(1997).  
 DR EMBL: AE001138; AAC66696.1;  
 DR TIGR: BB0314;

DR InterPro: IPR000092;  
 DR InterPro: IPR000235;  
 DR Pfam: PF00348; POLYPRENYL\_SYNTHET\_1;  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.

DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; UNKNOWN.1.  
 DR PROSITE: PS00052; RIBOSOMAL\_S7; UNKNOWN.1.  
 SQ SEQUENCE 347 AA; 39992 MW; EF51C634AC1F3DF CRC64;

Query Match 2.7%; Score 8; DB 2; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 IIDDIEDNS 70  
 Db 100 IIDDIEDNS 107

RESULT 12  
 O18869 PRELIMINARY; PRT; 606 AA.  
 AC O18869;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)

DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE SIMILARITY TO HUMAN LIPOPOLYSACCHARIDE-BINDING PROTEIN PRECURSOR.  
 GN C55C3.1.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Anscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,  
 RA Craixton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten T., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., McMurray B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.,  
 RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans."

RT Nature 368:32-38(1994).  
 RL [2]  
 RN SEQUENCE FROM N.A.

RP STRAIN-BRISTOL N2;  
 RC STRAIN-BRISTOL N2;  
 RA Woessle J., Stelljes L.,  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U53335; AAA96173.1;  
 DR InterPro: IPR001124;  
 DR Pfam: PF01273; LBP\_BP1\_CETP; 1.

DR SMART: SMO0329; BP12; 1.  
 SQ SEQUENCE 606 AA; 66957 MW; DD9982BF7940A6 CRC64;

Query Match 2.7%; Score 8; DB 5; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 153 GGLFGLAV 160  
 Db 191 GGLFGLAV 198

RESULT 13

O9JH67 PRELIMINARY; PRT; 844 AA.  
 AC O9JH67;

DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)  
 DE NONSTRUCTURAL PROTEIN, PUTATIVE SERINE PROTEASE.

OS sheep astrovirus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Astroviridae;  
 OC Astrovirus.  
 OX NCBI\_TaxID=70794;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-PROVIDED BY DR. D.R. SNOODGRASS, MOREDUN RESEARCH INSTITUTE,  
 RC EDINBURGH;  
 RX MEDLINE=98227920; PubMed=9568965;

RA Jonassen C.M., Jonassen T.O., Grinde B.;  
 RT "A common motif in the 3' end of the genomes of astroviruses, avian  
 RT infectious bronchitis virus and an equine rhinovirus."  
 RL J. Gen. Virol. 79:715-718(1998).  
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-PROVIDED BY DR. D.R. SNODGRASS, MOREDUN RESEARCH INSTITUTE,  
 RC EDINBURGH;  
 RA Jonassen C.M., Jonassen T., Saif Y., Snodgrass D., Ushijima H.,  
 RA Shimizu M., Grinde B.;  
 RT "Comparison of capsid sequences from human and animal astroviruses";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-PROVIDED BY DR. D.R. SNODGRASS, MOREDUN RESEARCH INSTITUTE,  
 RC EDINBURGH;  
 RA Jonassen C.M., Jonassen T., Sveen T., Snodgrass D., Grinde B.;  
 RT "Sequencing and molecular characterization of sheep astrovirus";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Y15937; CAB95002.1; -  
 SQ SEQUENCE 844 AA; 95454 MW; 4A8735C8C56ACB6C CRC64;

Query Match 2.7%; Score 8; DB 14; Length 844;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 140 TTEEKAM 147  
 |||||  
 Db 658 TTEEKAM 665

RESULT 14  
 007380 PRELIMINARY; PRT: 1790 AA.  
 AC Q07380; P89892;  
 DT 01-NOV-1996 (TREMblrel. 01, Created)  
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE HYPOHETICAL 206.5 KDA PROTEIN YD038W.  
 GN US01.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bloeker H., Brandt P.;  
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z74106; CAA98621.1; -  
 DR EMBL: Z74105; CAA98620.1; -  
 DR SGD: S0002216; US01.  
 DR InterPro: IPR002017;  
 SQ SEQUENCE 1790 AA; 206450 MW; 90062544F55A52EE CRC64;

Query Match 2.7%; Score 8; DB 3; Length 1790;  
 Best Local Similarity 100.0%; Pred. No. 48;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 173 LKPLNTL 180  
 |||||  
 Db 59 LKPLNTL 66

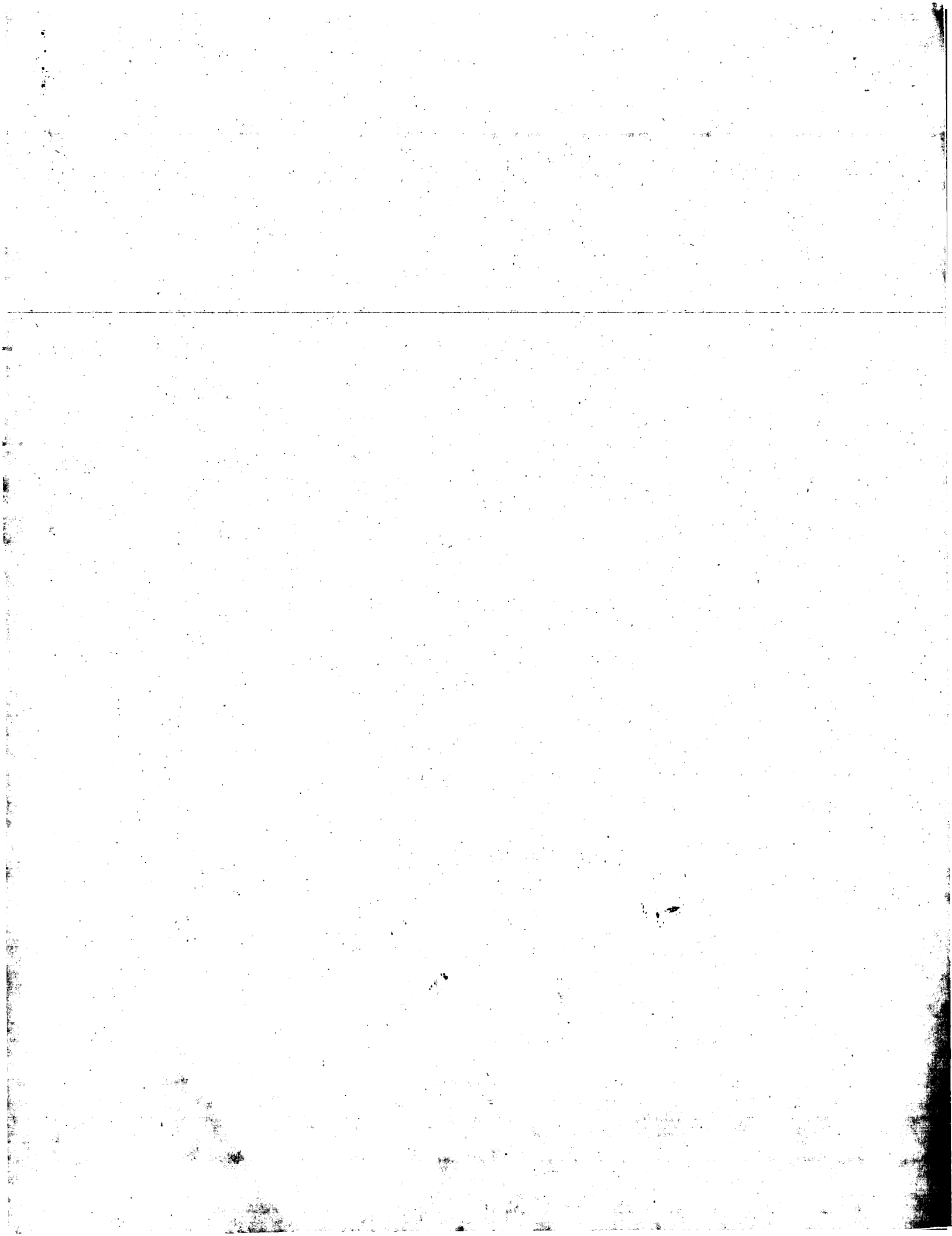
RESULT 15  
 09WS35 PRELIMINARY; PRT: 41 AA.  
 AC 09WS35;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)  
 DE GLYCOPROTEIN (FRAGMENT).  
 OS Rabies virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Rhabdoviridae; Lyssavirus.  
 OX NCBI\_TaxID=11292;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=M669;  
 RA Nel L.H., Jaftha J.B., von Teichmann B.F., Olivier M., Meredith C.;  
 RT "Vervet rabies in South Africa: A re-evaluation of the viruses  
 RT involved";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF079907; AAD39050.1; -  
 FT NON\_TER  
 SQ SEQUENCE 41 AA; 4569 MW; 5F2EA3855CDFDEE9 CRC64;

Query Match 2.3%; Score 7; DB 14; Length 41;  
 Best Local Similarity 100.0%; Pred. No. 18;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 223 SRPESTQ 229  
 |||||  
 Db 1 SRPESTQ 7

Search completed: August 19, 2001, 11:24:13  
 Job time: 122 sec







GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2001, 11:21:01 ; Search time 20.61 Seconds  
(without alignments) 882.445 Million cell updates/sec

Title: US-09-744-527-4

Perfect score: 300  
Sequence: 1 MEKQETVQRILLPEPKYL.....NPETVALVKHLSKMEKENE 300

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 412676 seqs, 60623988 residues

Word size : 0

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq\_0601:\*

1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT:\*  
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT:\*  
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT:\*  
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT:\*  
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT:\*  
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT:\*  
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT:\*  
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT:\*  
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT:\*  
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT:\*  
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT:\*  
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT:\*  
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT:\*  
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT:\*  
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT:\*  
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT:\*  
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT:\*  
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT:\*  
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT:\*  
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT:\*  
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT:\*  
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	300	100.0	300	21	AA68909
2	203	67.7	300	17	AA68909
3	203	67.7	300	17	AA68909
4	203	67.7	300	19	AA68909
5	203	67.7	300	20	AA68909
6	14	4.7	304	21	AA68909
7	7	2.3	376	18	AA68909
8	7	2.3	77	21	AA68909
9	7	2.3	77	21	AA68909
10	7	2.3	89	20	AA68909
11	7	2.3	111	21	AA68909
11	7	2.3	114	21	AA68909

12	7	2.3	119	19	AA68909
13	7	2.3	147	21	AA68909
14	7	2.3	152	21	AA68909
15	7	2.3	166	21	AA68909
16	7	2.3	166	21	AA68909
17	7	2.3	166	21	AA68909
18	7	2.3	166	21	AA68909
19	7	2.3	166	21	AA68909
20	7	2.3	166	21	AA68909
21	7	2.3	166	21	AA68909
22	7	2.3	166	21	AA68909
23	7	2.3	166	21	AA68909
24	7	2.3	166	21	AA68909
25	7	2.3	166	21	AA68909
26	7	2.3	166	21	AA68909
27	7	2.3	166	21	AA68909
28	7	2.3	166	21	AA68909
29	7	2.3	166	21	AA68909
30	7	2.3	166	21	AA68909
31	7	2.3	166	21	AA68909
32	7	2.3	166	21	AA68909
33	7	2.3	166	21	AA68909
34	7	2.3	166	21	AA68909
35	7	2.3	166	21	AA68909
36	7	2.3	166	21	AA68909
37	7	2.3	166	21	AA68909
38	7	2.3	166	21	AA68909
39	7	2.3	166	21	AA68909
40	7	2.3	166	21	AA68909
41	7	2.3	166	21	AA68909
42	7	2.3	166	21	AA68909
43	7	2.3	166	21	AA68909
44	7	2.3	166	21	AA68909
45	7	2.3	166	21	AA68909

#### ALIGNMENTS

RESULT 1	
AA68909	standard; Protein: 300 AA.
XX	
AC	AA68909:
XX	
DT	16-MAY-2000 (first entry)
XX	
DE	A geranylgeranyl pyrophosphate synthetase (hGPPS).
XX	
KW	Human: geranylgeranyl pyrophosphate synthetase; hGPPS; chromosome 1;
KW	1q42-1q43 locus; prostate cancer; hGPPS; diallelic marker;
KW	mevalonic biosynthetic pathway.
XX	
OS	Homo sapiens.
XX	
PN	W0200005382-A2.
XX	
PD	03-FEB-2000.
XX	
PF	23-JUL-1999: 99WO-IB01353.
XX	
PR	23-JUL-1998: 98US-0093940.
XX	
PA	(GEST) GENSET.
XX	
PI	Bougueret L;
XX	
DR	WPI: 2000-182704/15.
XX	
PT	N-PSDB: AA68909, AA68909.
PT	New isolated human geranylgeranyl pyrophosphate synthetase nucleic acids, used to develop agents for the diagnosis of, e.g. pathologies related to a defect in the mevalonic biosynthetic pathway

Pneumococcal sur  
Human pancreatic c  
HTRM clone 004586  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Human secreted pro  
Streptococcus pneu  
Arabidopsis thalia  
Arabidopsis thalia  
Arabidopsis thalia  
Corynebacterium gl  
Floral organ-speci  
R. gracilis D-amin  
Rhodospirillum tor  
Human secreted pro  
Corynebacterium gl  
80 kD protein. Te  
A serine/threonine  
Merizite apical-en  
Merizite apical-e  
Chimeric p53 link  
Tumour associated  
Prenyl diphosphate  
Immunogenic peptid  
HLA-A24 antigenic  
CD66 peptide CD66e  
HLA class II bind1  
CD66 peptide CD66a  
CD66 peptide CD66a  
CD66 peptide CD66d  
CD66 peptide CD66f  
HLA class II bind1

XX  
PS  
XX

Claim 15; Page 84-85; 88pp; English

The present sequence represents a human geranylgeranyl pyrophosphate synthetase (hGGPS). Two differently spliced mRNAs exist for this gene. The first spliced mRNA is derived from a cDNA (AAZ60888) which comprises exons 1, 2, 3 and 4. The second mRNA is derived from a cDNA (AAZ60889) which comprises 1b1s, 2, 3, and 4. The hGGPS gene is located on chromosome 1, at the 1q42-1q43 locus. This chromosome 1 locus has been shown to carry a predisposing gene for prostate cancer. The nucleic acids encoding hGGPS can be used for screening for agents which modulate the expression of the hGGPS gene. Such agents can be used in therapeutic applications. The diallelic markers associated with the hGGPS gene can be used for the diagnosis of diseases related to an alteration in the regulatory or coding regions of hGGPS, such as pathologies related to a defect in the mevalonic biosynthetic pathway. The products can also be used for detection, diagnosis and drug screening.

Sequence 300 AA;

Query Match	100.0%	Score:300;	DB 21;	Length 300;
Best Local Similarity	100.0%	Pred. No. 4	4e-297;	
Matches 300;	Conservative 0;	Mismatches 0;	Totals 0;	Coverage 0

QY	1	MEKHOEYVORLLEPEPKYLLIOLPGQOVRIKKSQAFNMELKVPEDKLOLIITEVTEMLHNAS	60
Db	1	mekqetvqglllepkyllilpgqkvrtklsqatnmlkvpedkllilevtemlhnas	60
QY	61	LLIDDDIEDNSKLRGPPVAHSITGIPSVINSANYEFGLEKVTLDHPDAVLEFTRQL	12
Db	61	lllddiednslrrgppvashisylgipsvinsanyyflgklevtlldhpavvlftqrll	12
QY	121	ELHOGGGLIDYMRDNTCTEETEEYFAAMLOKTGGFLGAVGIMQLDSVDKEDKLPMLNTL	180
Db	121	elbhggglidylvrndytcpteeeyfayamlqktcggllfglavglmqltsdykedkplmlntl	180
QY	181	GLFFQIIRDDVAMNHSKEYSENKSFEDDTECKSEFPPIHAIWNRPESTOVONILQRTEN	240
Db	181	glffqirddvamnhskeysenksfceddteqkfsifpilhawnrpestqvpnilqrten	240
QY	241	IDIKKYCVHLEDVGSFEYTRNTLKELEAKKAYKOIDRGNGPELTVLWYHNSKMKKEENE	300
Db	241	idikkycvmyledvgsfeytrntlkeleakaykgidargngpelvalvwhnskmlkkeene	300

## RESULT 2

AAK9/565 standard; Protein; 300 AA.

AC . AAR97565;

DT 17-FEB-1997 (first entry)  
XX

geranylgeranyl pyrophosphate synthetase (GGPPS).

KW steroid; post-translational modification; gene therapy;  
 rhabdomyolysis; cell morphology; neoplasia; cancer.  
 XX

**Homo sapiens.**

FT	key	Location/Qualifiers
75	Region	61 75

Region

Region	180	183
/note= "conserved aspartate motif"		

ET

```

/1101e= "conserved aspartate motif"
XX
XX

```

PN W09621736-A1

PD 18-JUL-1996

XX

PF 11-JAN-1995; 95WO-US00421  
XX

PR 11-JAN-1995; 95WO-US00421.  
XX

PA (HUMA-) HUMAN GENOME SCI INC  
XX

XX  
FI  
greene JM, Kirkness EF, Rosen CA;

DK WPI; 1996-342289/34  
DR N-PSDB; AAT31846

DK N-PSDB; AAT31846.  
YY

PT1 polynucleotide encoding human geranyl:geranyl pyrophosphate synthetase - fragments of which can be used as probes to diagnose disease or susceptibility to disease related to mutation in above polynucleotide

PS Claim 5; page 39-40; 55pp; English.  
XX

The mouse cDNA sequence encoding GGPPS was obtained from a foetal heart cDNA library and identified by homology to the corresponding enzyme of *Neurospora crassa*. The GGPS enzyme is involved in post-translational modification (more specifically the geranylgeranylation) of proteins and is also involved in a branch of the cholesterol/steroid metabolic pathway. GGPS catalyses the transfer of 3 molecules of isopentenyl diphosphate onto dimethylallyl diphosphate to form C<sub>60</sub> GGPP. GGPP controls the signal transduction, activation and trafficking of GGGP-modified proteins. Human GGPS can be used to control cell morphology [e.g., to treat rhabdomyosarcoma], side effect of treatment with HMG-CoA reductase), and unwanted apoptosis. Antagonists of GGPS can be used to treat neoplasms and to prevent viral proliferation. Increased levels of GGPS are indicative of cancer cell growth.

Sequence 300 AA;

Query Match	67.7%	Score 203	DB 17	Length 300
Best Local Similarity	100.0%	Pred. No. 2.4e+198		
Matches 203; Conservative	0	Mismatches 0	Indels 0	Cases 0

[illegible]

### RESULT 3

AAW59753 standard; Protein; 300 AA

AAW59753;

12-OCT-1998 (first entry)

Human geranyl geranyl pyrophosphate synthase

agonist; ss. pyrophosphate synthase; hggps; antagonist,

**Homo sapiens**

**X**

PN US5786193-A.  
 XX 28-JUL-1998.  
 PD 06-JUN-1995; 95US-0469665.  
 XX 06-JUN-1995; 95US-0469665.  
 PR 06-JUN-1995; 95US-0469665.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Greene JM, Kirkness EF, Rosen CA;  
 DR WPI; 1998-436529/37.  
 DR N-PSDB; AAV41555.  
 XX  
 PT DNA encoding human geranylgeranyl pyrophosphate synthetase  
 PT polypeptide - useful for producing recombinant polypeptide  
 PS  
 PS Claim 1; Fig 1; 22pp; English.  
 CC This is the amino acid sequence of the human geranyl geranyl  
 CC pyrophosphate synthase (hGPPS) used in the method of the invention.  
 CC The proteins are used in the production of recombinant hGPPS, which  
 CC can be used to screen for hGPPS antagonists or agonists.  
 XX  
 SQ Sequence 300 AA;

Query Match 67.7%; Score 203; DB 19; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-198;  
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKTEVORILLEPPYVYLQLPKQVRTKLSQAFNHLKVPEDKLIIEVTEMLNAS 60  
 Db 1 MEKTEVGRILLLEPYVYLIQIPGKYRTKLSQAFNHLKVPEDKLIIEVTEMLNAS 60  
 QY 61 LLIDIEDNSKLRGFPVANSIYIGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROL 120  
 Db 61 LLIDIEDNSKLRGFPVANSIYIGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROL 120  
 QY 121 ELHOGGLDIYWRDNYCPTPEEYKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
 Db 121 ELHOGGLDIYWRDNYCPTPEEYKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
 QY 181 GLFPIRDYANLHSEYSENKS 203  
 Db 181 GLFPIRDYANLHSEYSENKS 203

RESULT 4  
 ID AAY27005 standard; Protein; 300 AA.  
 AC AAY27005;

XX 20-SEP-1999 (first entry)  
 DE Human geranylgeranyl pyrophosphate synthetase (hGPPS) polypeptide.  
 DE Geranylgeranyl pyrophosphate synthetase; hGPPS; cell morphology; human;  
 KW apoptosis; neoplasia; tumour; cancer; viral proliferation; gene therapy;  
 KW human choroideremia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5928924-A.  
 PD 27-JUL-1999.  
 XX  
 PF 11-MAR-1998; 98US-0038596.  
 XX 06-JUN-1995; 95US-0469665.  
 PR 11-JAN-1995; 95WO-US00421.  
 PA

PR 11-MAR-1998; 98US-0038596.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Greene JM, Kirkness EF, Rosen CA;  
 DR WPI; 1999-429504/36.  
 DR N-PSDB; AAX86773.  
 XX  
 PT Geranylgeranyl pyrophosphate synthetase peptides useful for  
 PT regulating apoptosis  
 PS  
 PS Claim 1; Fig 1A-D; 23pp; English.  
 CC This represents a human geranylgeranyl pyrophosphate synthetase (hGPPS)  
 CC polypeptide. The hGPPS polynucleotides and polypeptides may be used to  
 CC control the morphology of cells, suppress apoptosis, screen for hGPPS  
 CC agonists and antagonists (which may in turn be used to treat neoplasia,  
 CC e.g. tumours and cancer cell growth, and prevent viral proliferation),  
 CC gene therapy, and diagnose a disease or susceptibility to diseases  
 CC related to a mutation in hGPPS (e.g. human choroideremia).  
 XX  
 SQ Sequence 300 AA;

Query Match 67.7%; Score 203; DB 20; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-198;  
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKTEVORILLEPPYVYLQLPKQVRTKLSQAFNHLKVPEDKLIIEVTEMLNAS 60  
 Db 1 MEKTEVGRILLLEPYVYLIQIPGKYRTKLSQAFNHLKVPEDKLIIEVTEMLNAS 60  
 QY 61 LLIDIEDNSKLRGFPVANSIYIGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROL 120  
 Db 61 LLIDIEDNSKLRGFPVANSIYIGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROL 120  
 QY 121 ELHOGGLDIYWRDNYCPTPEEYKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
 Db 121 ELHOGGLDIYWRDNYCPTPEEYKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
 QY 181 GLFPIRDYANLHSEYSENKS 203  
 Db 181 GLFPIRDYANLHSEYSENKS 203

RESULT 5  
 ID AAB58276 standard; Protein; 304 AA.  
 AC AAB58276;

XX 14-MAR-2001 (first entry)  
 DE Lung cancer associated polypeptide sequence SEQ ID 614.  
 DE Human; lung cancer associated protein; neuroprotective; cytostatic;  
 KW cardioactive; immunomodulatory; muscular active; vulnerrary;  
 KW gastrointestinal; nephrotoxic; antiinfective; gynecological;  
 KW antibacterial; diagnosis; neural disorder; immune disorder; reproductive;  
 KW proliferative disorder; wound healing; infectious disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200055180-A2.  
 PD 21-SEP-2000.  
 XX  
 PF 08-MAR-2000; 2000MO-US05918.  
 XX 12-MAR-1999; 99US-0124270.  
 PR  
 PA (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.  
 PI Ruben SM;  
 XX WPI: 2000-587514/55.  
 XX N-PSDB: AAF18152.  
 DR  
 XX Lung cancer associated gene sequences, referred to as lung cancer  
 PT antigens, useful for treatment, prevention, and diagnosis of disorders  
 PT such as lung cancer -  
 XX  
 PS Claim 11; Page 1108-1109; 1425pp; English.  
 XX  
 CC Polynucleotide sequences AAF1982 - AAF18424 encode human lung cancer  
 CC associated proteins represented in AAB58106 - AAB58548. Lung cancer  
 CC associated proteins and polynucleotide sequences, their agonists, and  
 CC antagonists may have neuroprotective, cytostatic, cardioactive, and  
 CC immunomodulatory; muscular active general; vulnarary; gastrointestinal  
 CC activity; nephrotropic; antineoplastic; gynecological; or antibacterial  
 CC protein or polynucleotide sequences. The lung cancer associated  
 CC polynucleotide sequences may be used for detection of lung cancer,  
 CC chromosome identification, as chromosome markers, and for numerous other  
 CC diagnostic or research purposes. The proteins may be used to treat  
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative  
 CC disorders. The proteins may also be used in the treatment of wounds and  
 CC infectious diseases. Polynucleotide sequences AAF18425 - AAF18433 and  
 CC peptide AAB58549 are used in the course of the invention for the  
 CC identification and characterisation of the polynucleotide and protein  
 CC sequences.  
 XX  
 SQ Sequence 304 AA:  
 Query Match 67.7%; Score 203; DB 21; Length 304;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-198;  
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MERTQETVORILLEPKYVLLQPGKQVRLKSOANHWLKVPEKLIITIEVTEMLNANS 60  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 5 mektqetvtrlllepkyllqllpkyvrtklsgatnhtkvpedklllilevtemlnans 64  
 QY 61 LIIDIEDNSKLRGPPVAHSIYGIPIVINSANYVFLGLEKVLTLDPDAVKFTQGL 120  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 65 liiddednsklrrgfpvshsiygiipivinsanyvflglekvltdpdaavkllftqgl 124  
 QY 121 EHHOGAGIDIVKNDYTCPTEEETKAMVLOKTGIFGLAVGIMOLFSDYKEDLPLINTL 180  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 125 elhggagldivrtndytcpteeeykamvlgktgglfglavglmglfedykedlplintl 184  
 QY 181 GTFQIRDDYANLHSEKSEKNS 203  
 Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 185 gltfqirddyanlhseksenks 207  
 RESULT 6  
 ID AAW22498 standard; Protein: 376 AA.  
 XX  
 AC AAW22498;  
 XX  
 DT 10-MAR-1998 (first entry).  
 XX  
 DE Phaffia derived carotenoid biosynthesis pathway enzyme PcrCtE.  
 XX  
 KW Glyceraldehyde-3-phosphate dehydrogenase; GAPDH; carotenoid;  
 KW synthesis; promoter; recombinant DNA; astaxanthin; ribosomal protein;  
 KW food colouring.  
 XX  
 OS Phaffia rhodozyma.  
 XX  
 PN MO9723633-A1.  
 XX

XX 03-JUL-1997.  
 PD 23-DEC-1996; 96WO-EP05887.  
 XX  
 PF 11-APR-1996; 96EP-0200943.  
 PR 22-DEC-1995; 95EP-0203620.  
 XX  
 PA (KONN) GIST-BROCADES BV.  
 PA (COOLU) COOLUEN A J J.  
 XX  
 PI Verdoes JC, Wery J;  
 DR WPI: 1997-351068/32.  
 DR N-PSDB: AAT72942.  
 XX  
 PT Phaffia derived GAPDH and carotenoid synthesis genes and promoter  
 PT fragment - used in the recombinant production of therapeutically  
 PT useful proteins e.g. carotenoids for use in food colouring  
 XX  
 PS Claim 11; Page 55-56; 118pp; English.  
 XX  
 CC The present sequence represents a Phaffia derived carotenoid  
 CC biosynthesis pathway enzyme. The nucleic acid encoding this protein  
 CC can be used in the novel recombinant DNA of the present invention. The  
 CC recombinant DNA comprises a transcription promoter operably linked to a  
 CC downstream sequence to be expressed, where the transcription promoter  
 CC comprises a region found upstream of the open reading frame (ORF) of a  
 CC highly expressed Phaffia gene (preferably GAPDH, a ribosomal protein  
 CC gene, an enzyme involved in the biosynthesis pathway). The recombinant  
 CC DNA can be used to transform hosts, preferably Phaffia. These  
 CC transformed hosts are then used in the recombinant production of GAPDH  
 CC or an enzyme involved in carotenoid synthesis, preferably astaxanthin.  
 CC They may also be used to produce a pharmaceutical product. Purified  
 CC carotenoids can be used as colourants in food and/or feed, and also in  
 CC cosmetics.  
 XX  
 SQ Sequence 376 AA:  
 Query Match 4.7%; Score 14; DB 18; Length 376;  
 Best Local Similarity 100.0%; Pred. No. 7e-06;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 206 EDLTEGKFSPTTH 219  
 Db ||||||||||||||||  
 263 edltegkfsptth 276  
 RESULT 7  
 ID AAB29887 standard; Protein: 77 AA.  
 XX  
 AC AAB29887;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein BLAST search protein SEQ ID NO: 145.  
 XX  
 KW Cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;  
 KW vulnarary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200061779-A1.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 06-APR-2000; 2000WO-US09068.  
 XX

PR 09-APR-1999; 99US-0128699.  
 PR 20-JAN-2000; 2000US-0177050.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX WPI; 2000-647424/62.  
 DR  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 PT  
 XX Disclosure; Page 472; 495pp; English.  
 PS  
 XX The invention relates to the isolation of genes AAC63410-C63458 encoding  
 CC the human secreted proteins AAB29802-B29850. This sequence represents a  
 CC peptide fragment homologous to the protein encoded by the gene given  
 CC in the descriptor line. The sequence is a search result from a BLASTX  
 CC homology search. The genes and proteins are useful for preventing,  
 CC ameliorating or treating medical conditions, e.g. by protein or gene  
 CC therapy. The genes are isolated from a range of human tissues disclosed  
 CC in the specification. The nucleic acids, proteins, antibodies and  
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such  
 CC as myocardial ischaemias; (d) wound healing; (e) neurological diseases  
 CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such  
 CC as viral, bacterial, fungal and parasitic infections.  
 XX  
 SQ Sequence 77 AA:

Query Match 2.3%; Score 7; DB 21; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 278 RGNPEL 284  
 |||||  
 Db 15 rgnpel 21

RESULT 8  
 AAB29888 standard; Protein: 77 AA.  
 XX  
 AC AAB29888;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein BLAST search protein SEQ ID NO: 146.  
 XX  
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;  
 KM valnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KM neurological disease; infection; human; secreted protein.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200061779-A1.  
 XX  
 PD 19-OCT-2000.  
 XX  
 PF 06-APR-2000; 2000MO-US09068.  
 XX  
 PR 09-APR-1999; 99US-0128699.  
 XX 20-JAN-2000; 2000US-0177050.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.

XX  
 PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX WPI; 2000-647424/62.  
 DR  
 XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -  
 PT  
 XX Disclosure; Page 472; 495pp; English.  
 PS  
 XX The invention relates to the isolation of genes AAC63410-C63458 encoding  
 CC the human secreted proteins AAB29802-B29850. This sequence represents a  
 CC fragment of the protein encoded by the gene given in the descriptor  
 CC line. The sequence is used as a query sequence for doing BLASTX searches  
 CC to determine homologous sequence to the protein. The genes and proteins  
 CC are useful for preventing, ameliorating or treating medical conditions,  
 CC e.g. by protein or gene therapy. The genes are isolated from a range of  
 CC human tissues disclosed in the specification. The nucleic acids,  
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,  
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,  
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders  
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis; (c)  
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and  
 CC (f) infectious diseases such as viral, bacterial, fungal and parasitic  
 XX infections.  
 XX  
 SQ Sequence 77 AA:

Query Match 2.3%; Score 7; DB 21; Length 77;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 278 RGNPEL 284  
 |||||  
 Db 15 rgnpel 21

RESULT 9  
 AAY34930 standard; Protein: 89 AA.  
 XX  
 AC AAY34930;  
 XX  
 DT 13-SEP-1999 (first entry)  
 XX  
 DE Chlamydia pneumoniae ribosomal polypeptide.  
 XX  
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;  
 KM sinusitis; purulent otitis media; erythema nodosum; pharyngitis;  
 KM vaccine; neutralising epitope.  
 XX  
 OS Chlamydia pneumoniae.  
 XX  
 PN WO9927105-A2.  
 XX  
 PD 03-JUN-1999.  
 XX  
 PF 20-NOV-1998; 98MO-IB01890.  
 XX  
 PR 04-NOV-1998; 98US-0107078.  
 XX 21-NOV-1997; 97FR-0014673.  
 XX  
 PA (GENST ) GENSET.  
 XX  
 PI Gliffais R;  
 XX WPI; 1999-357842/30.  
 XX

PT Genome sequence of Chlamydia pneumoniae  
 XX Page 857; Disclosure: 1912pp; English.  
 XX  
 CC AAY34584-Y35879 represent the proteins encoded by all the open reading  
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae.  
 CC C. pneumoniae causes respiratory disease such as pneumonia and  
 CC bronchitis and is thought to be a contributing factor in heart  
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema  
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading  
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in  
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae  
 CC nucleotides sequences can also be used as immunogenic compositions,  
 CC especially where the vector directs the expression of a neutralising  
 CC epitope of C. pneumoniae.  
 CC  
 XX Sequence 89 AA:  
 SQ

Query Match Best Local Similarity 2.38; Score 7; DB 20; Length 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 LGLEKVL 104  
 |||||  
 Db 75 LGLEKVL 81

RESULT 10  
 AAG13400  
 ID AAG13400 standard; protein, 111 AA.  
 XX  
 AC AAG13400;

DT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 12882;

DE Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KM termination sequence.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129645.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130691.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 07-MAY-1999; 99US-0132486.  
 PR 11-MAY-1999; 99US-0132863.  
 PR 99US-0134256.

PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 14-MAY-1999; 99US-0134370.  
 PR 18-MAY-1999; 99US-0134766.  
 PR 19-MAY-1999; 99US-0134941.  
 PR 20-MAY-1999; 99US-0135124.  
 PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137528.  
 PR 08-JUN-1999; 99US-0137724.  
 PR 10-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139119.  
 PR 17-JUN-1999; 99US-0139452.  
 PR 18-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 28-JUN-1999; 99US-0140695.  
 PR 29-JUN-1999; 99US-0140823.  
 PR 30-JUN-1999; 99US-0140991.  
 PR 01-JUL-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142055.  
 PR 08-JUL-1999; 99US-0142390.  
 PR 09-JUL-1999; 99US-0142803.  
 PR 12-JUL-1999; 99US-0142920.  
 PR 13-JUL-1999; 99US-0142977.  
 PR 14-JUL-1999; 99US-0143542.  
 PR 15-JUL-1999; 99US-0143624.  
 PR 16-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 20-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 23-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.

PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 2.3%; Score 7; DB 21; Length 111;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 YFLGLEK 102.  
Db |||||  
22 yflglek 28

RESULT 11  
AAG13399  
ID AAG13399 standard; Protein; 114 AA.  
XX  
AC AAG13399;  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 12881.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
OS Arabidopsis thaliana.  
XX  
PN EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX

PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.  
 PR 24-MAY-1999; 99US-0135629.  
 PR 25-MAY-1999; 99US-0136021.  
 PR 27-MAY-1999; 99US-0136392.  
 PR 28-MAY-1999; 99US-0136782.  
 PR 01-JUN-1999; 99US-0137222.  
 PR 03-JUN-1999; 99US-0137528.  
 PR 04-JUN-1999; 99US-0137724.  
 PR 07-JUN-1999; 99US-0138094.  
 PR 08-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0139839.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140655.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142205.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142820.  
 PR 12-JUL-1999; 99US-0142877.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 15-JUL-1999; 99US-0143624.  
 PR 16-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 19-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 09-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 10-AUG-1999; 99US-0147935.  
 PR 11-AUG-1999; 99US-0148171.  
 PR 12-AUG-1999; 99US-0148319.  
 PR 13-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 16-AUG-1999; 99US-0148684.  
 PR 17-AUG-1999; 99US-0149368.  
 PR 18-AUG-1999; 99US-0149175.  
 PR 20-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 23-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 25-AUG-1999; 99US-0149930.  
 PR 26-AUG-1999; 99US-0150566.  
 PR 27-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151438.  
 PR 07-SEP-1999; 99US-0151930.  
 PR 10-SEP-1999; 99US-0152363.  
 PR 13-SEP-1999; 99US-0153070.  
 PR 15-SEP-1999; 99US-0153758.  
 PR 16-SEP-1999; 99US-0154018.  
 PR 20-SEP-1999; 99US-0154039.  
 PR 22-SEP-1999; 99US-0154779.  
 PR 23-SEP-1999; 99US-0155139.  
 PR 24-SEP-1999; 99US-0155486.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 29-SEP-1999; 99US-0156458.  
 PR 04-OCT-1999; 99US-0156596.  
 PR 05-OCT-1999; 99US-0157117.  
 PR 06-OCT-1999; 99US-0157753.  
 PR 07-OCT-1999; 99US-0157865.  
 PR 08-OCT-1999; 99US-0158029.  
 PR 12-OCT-1999; 99US-0158232.  
 PR 13-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159284.  
 PR 13-OCT-1999; 99US-0159285.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 18-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159684.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 22-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 23-OCT-1999; 99US-0160981.  
 PR 25-OCT-1999; 99US-0160989.  
 PR 25-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 26-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.





ID AAY73326 standard; Protein; 152 AA.  
 XX AAY73326;  
 AC AAY73326;  
 XX  
 DT 24-FEB-2000 (first entry)  
 XX  
 DE HTRM clone 004586 protein sequence  
 XX  
 XX HTRM; human transcriptional regulatory molecule; arteriosclerosis; AIDS;  
 KW arteriosclerosis; cirrhosis; cancer; leukemia; diabetes mellitus;  
 KW Addison's disease; multiple sclerosis; rheumatoid arthritis; infection;  
 KW trauma; myasthenia gravis; adenocarcinoma; immune disorder; treatment.  
 XX  
 OS Homo sapiens.  
 XX  
 XX -W0957144-A2.  
 PN 11-NOV-1999.  
 XX  
 PD 04-MAY-1999; 99WO-US09935.  
 XX  
 PE 05-MAY-1998; 98US-0084254.  
 PR 07-AUG-1998; 98US-0095827.  
 PR 02-OCT-1998; 98US-0102745.  
 XX  
 PA (INCYTE) INCYTE PHARM INC.  
 XX  
 XX Hillman JL, Bandman O, Lal P, Yue H, Reddy R, Tang YT;  
 PI Gerstin EH, Patterson C, Baughn MR, Azimzai Y, Lu DAM;  
 XX  
 DR WPI: 2000-052941/04.  
 DR N-PSDB; AAZ52411.  
 XX  
 PT New peptides useful for diagnosis, prevention and treatment of cancer  
 PT and immune disorders  
 PS  
 PS Claim 1: Page 90; 193pp; English.  
 XX  
 CC AAY73325-Y73389 are human transcriptional regulator molecule (HTRM)  
 CC protein sequences. The HTRM protein and nucleotide sequences are useful  
 CC for preventing or treating disorders associated with decreased expression  
 CC or activity of HTRM which include cell proliferative disorders such as  
 CC arteriosclerosis and cirrhosis; cancers including adenocarcinoma and  
 CC leukemia; immune disorders such as AIDS, Addison's disease, diabetes  
 CC mellitus, rheumatoid arthritis, multiple sclerosis, systemic lupus  
 CC erythematosus, and myasthenia gravis; infections and trauma. Antagonists  
 CC of the HTRM polypeptides are useful for treating or preventing disorders  
 CC associated with increased expression or activity of HTRMs. HTRM  
 CC polypeptides, their immunogenic fragments or oligopeptides are useful  
 CC for screening libraries of compounds in drug screening techniques.  
 CC of mRNA and regulating gene function by blocking the transcription  
 CC HTRM. Vectors expressing HTRM or agonists can also be used to prevent or  
 CC treat disorder associated with decreased HTRM expression. Antibodies  
 CC which specifically bind HTRM and polynucleotides encoding HTRM are useful  
 CC for diagnosing disorders associated with the expression of HTRM,  
 CC particularly in assays that detect the expression of HTRM. Nucleotide  
 CC sequences encoding HTRM may be useful to generate hybridization probes  
 CC useful in mapping the naturally occurring genomic sequence and to detect  
 CC differences in gene sequences among normal, carrier and affected  
 CC individuals. Using diagnostic assays, cancer can be detected prior to  
 CC the appearance of clinical symptoms and thereby progression of cancer can  
 CC be prevented by aggressive treatment or preventive measures.  
 CC  
 CC Sequence 152 AA;  
 XX  
 XX

Query Match 2.38; Score 7; DB 21; Length 152;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 263 TLKLEEA 269  
 |||||||

Db 83 tlkleea 89  
 RESULT 15  
 ID AAG15978  
 XX AAG15978 standard; Protein; 166 AA.  
 XX  
 AC AAG15978;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 16441.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX EP1033405-A2.  
 PN 06-SEP-2000.  
 PD  
 PD 25-FEB-2000; 2000EP-0301439.  
 PF  
 XX 25-FEB-1999; 99US-0121825.  
 XX 05-MAR-1999; 99US-0123180.  
 XX 09-MAR-1999; 99US-0123548.  
 XX 23-MAR-1999; 99US-0125788.  
 XX 25-MAR-1999; 99US-0126284.  
 XX 29-MAR-1999; 99US-0126785.  
 XX 01-APR-1999; 99US-0127462.  
 XX 06-APR-1999; 99US-0128234.  
 XX 08-APR-1999; 99US-0128714.  
 XX 16-APR-1999; 99US-0129845.  
 XX 19-APR-1999; 99US-0130077.  
 XX 21-APR-1999; 99US-0130449.  
 XX 23-APR-1999; 99US-0130510.  
 XX 23-APR-1999; 99US-0130891.  
 XX 28-APR-1999; 99US-0131449.  
 XX 30-APR-1999; 99US-0132048.  
 XX 30-APR-1999; 99US-0132407.  
 XX 04-MAY-1999; 99US-0132484.  
 XX 05-MAY-1999; 99US-0132485.  
 XX 06-MAY-1999; 99US-0132486.  
 XX 06-MAY-1999; 99US-0132487.  
 XX 07-MAY-1999; 99US-0132863.  
 XX 11-MAY-1999; 99US-0134256.  
 XX 14-MAY-1999; 99US-0134218.  
 XX 14-MAY-1999; 99US-0134219.  
 XX 14-MAY-1999; 99US-0134221.  
 XX 18-MAY-1999; 99US-0134370.  
 XX 19-MAY-1999; 99US-0134768.  
 XX 20-MAY-1999; 99US-0134941.  
 XX 21-MAY-1999; 99US-0135124.  
 XX 21-MAY-1999; 99US-0135353.  
 XX 24-MAY-1999; 99US-0135629.  
 XX 25-MAY-1999; 99US-0136021.  
 XX 27-MAY-1999; 99US-0136392.  
 XX 28-MAY-1999; 99US-0136782.  
 XX 01-JUN-1999; 99US-0137222.  
 XX 01-JUN-1999; 99US-0137528.  
 XX 03-JUN-1999; 99US-0137502.  
 XX 04-JUN-1999; 99US-0137724.  
 XX 07-JUN-1999; 99US-0138094.  
 XX 08-JUN-1999; 99US-0138340.  
 XX 10-JUN-1999; 99US-0138847.  
 XX 14-JUN-1999; 99US-0139119.  
 XX 16-JUN-1999; 99US-0139452.  
 XX 16-JUN-1999; 99US-0139453.  
 XX 17-JUN-1999; 99US-0139454.  
 XX 18-JUN-1999; 99US-0139455.  
 XX 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139751.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 28-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0142977.  
 PR 13-JUL-1999; 99US-0143542.  
 PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 20-JUL-1999; 99US-0144684.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145085.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 28-JUL-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 25-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158222.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 14-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0160989.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 25-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 26-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 28-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161992.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 2.38; Score 7; DB 21; Length 166;  
 Best Local Similarity 100.0%; Pred. No. 43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 96 yfLEK 102  
 |||||  
 Db 77 yfLEK 83

Search completed: August 19, 2001, 11:23:01  
 Job time: 120 sec





1

agonist; ss.

KW Human; geranyl geranyl pyrophosphate synthase; hGGPS; antagonist;

Human; geranyl pyrophosphate synthase; hGPPS; antagonist;

XX OS Homo sapiens.  
XX PN US5786193-A.  
XX 28-JUL-1998.  
XX PD 06-JUN-1995; 95US-0469665.  
XX PF 06-JUN-1995; 95US-0469665.  
XX PR 06-JUN-1995; 95US-0469665.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Greene JM, Kirkness EF, Rosen CA;  
XX DR WPI: 1998-436529/37.  
XX DR N-PSDB: AAV41555.  
XX PT DNA encoding human geranylgeranyl pyrophosphate synthetase  
XX PS polypeptide - useful for producing recombinant polypeptide.  
XX PS Claim 1; Fig 1; 22pp; English.  
XX CC This is the amino acid sequence of the human geranyl geranyl  
XX CC pyrophosphate synthase (hgpps) used in the method of the invention.  
XX CC The proteins are used in the production of recombinant hgpps, which  
XX CC can be used to screen for hgpps antagonists or agonists.  
XX SQ Sequence 300 AA;

Query Match 97.3%; Score 1523; DB 19; Length 300;  
Best Local Similarity 98.3%; Pred. No. 6.7e-139;  
Matches 295; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MEKTEVQRIILEPYKYLQIPGKQVTKLSQAFNHMLKVPEDKIQIIEVTEMLHNAS 60  
DB 1 mektetvqrilllepykylqipgkqvrtklsqafnhmlkvpedkiiqilevtemlhnas 60  
QY 61 LIIDIEDNSKLRGFPVAHSIYGIPSVINSANYYFLGLEKVLTLDPDAVKLFTROL 120  
DB 61 lliiddiednsklrgfpvashsiygi psvinsanyyflglekvltlldpdkvklftroql 120  
QY 121 ELHOGGGLDIWRDNYTCPTTEEEYKAMVLQKTGGLFGLAVGLMQLFSYKEDLKPDLNTL 180  
DB 121 elhoggglidiwrndytcptteeykamlv lqktgglf gllavglmq lfsdykedl kpldntl 180  
QY 181 GLFQIRDDYANLHSEKSEKSFCEDLTEGFSPTTHATMSRESNOVONITLQRTEN 240  
DB 181 glfqi rddyanlhseksenk sfcedltegfs ptt hatisresno vovonitl qrt en 240  
QY 241 IDIKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNDELVALVKHLSKMKKEENE 300  
DB 241 idikkycvhyledvgsgeytrntlkleakaykqidargnpe lvalvkhlskmskeene 300

RESULT 4  
AAV27005  
ID AAV27005 standard; Protein; 300 AA.  
AC AAV27005;  
XX 20-SEP-1999 (first entry).  
XX Human geranylgeranyl pyrophosphate synthetase (hgpps) polypeptide.  
XX Geranylgeranyl pyrophosphate synthetase; hgpps; cell morphology; human;  
XX apoptosis; neoplasia; tumour; cancer; viral proliferation; gene therapy;  
XX human choroidermlna.  
XX Homo sapiens.  
XX OS  
XX PN US5928924-A.

XX OS 27-JUL-1999.  
XX PD 11-MAR-1998; 98US-0038596.  
XX PF 06-JUN-1995; 95US-0469665.  
XX PR 11-JAN-1995; 95MO-US00421.  
XX PR 11-MAR-1998; 98US-0038596.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Greene JM, Kirkness EF, Rosen CA;  
XX DR WPI: 1999-429504/36.  
XX DR N-PSDB: AAX86773.  
XX PT Geranylgeranyl pyrophosphate synthetase peptides useful for  
XX PS regulating apoptosis  
XX PS Claim 1; Fig 1A-D; 23pp; English.  
XX CC This represents a human geranylgeranyl pyrophosphate synthetase (hgpps)  
XX CC polypeptide. The hgpps polynucleotides and polypeptides may be used to  
XX CC control the morphology of cells, suppress apoptosis, screen for hgpps  
XX CC agonists and antagonists (which may in turn be used to treat neoplasia,  
XX CC e.g. tumours and cancer cell growth, and prevent viral proliferation),  
XX CC gene therapy, and diagnose a disease or susceptibility to diseases  
XX CC related to a mutation in hgpps (e.g. human choroidermlna).  
XX SQ Sequence 300 AA;

Query Match 97.3%; Score 1523; DB 20; Length 300;  
Best Local Similarity 98.3%; Pred. No. 6.7e-139;  
Matches 295; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MEKTEVQRIILEPYKYLQIPGKQVTKLSQAFNHMLKVPEDKIQIIEVTEMLHNAS 60  
DB 1 mektetvqrilllepykylqipgkqvrtklsqafnhmlkvpedkiiqilevtemlhnas 60  
QY 61 LIIDIEDNSKLRGFPVAHSIYGIPSVINSANYYFLGLEKVLTLDPDAVKLFTROL 120  
DB 61 lliiddiednsklrgfpvashsiygi psvinsanyyflglekvltlldpdkvklftroql 120  
QY 121 ELHOGGGLDIWRDNYTCPTTEEEYKAMVLQKTGGLFGLAVGLMQLFSYKEDLKPDLNTL 180  
DB 121 elhoggglidiwrndytcptteeykamlv lqktgglf gllavglmq lfsdykedl kpldntl 180  
QY 181 GLFQIRDDYANLHSEKSEKSFCEDLTEGFSPTTHATMSRESNOVONITLQRTEN 240  
DB 181 glfqi rddyanlhseksenk sfcedltegfs ptt hatisresno vovonitl qrt en 240  
QY 241 IDIKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNDELVALVKHLSKMKKEENE 300  
DB 241 idikkycvhyledvgsgeytrntlkleakaykqidargnpe lvalvkhlskmskeene 300

RESULT 5  
AAR97565  
ID AAR97565 standard; Protein; 300 AA.  
AC AAR97565;  
XX 17-FEB-1997 (first entry)  
XX Human geranylgeranyl pyrophosphate synthetase (Ggpps).  
XX Geranylgeranyl pyrophosphate synthetase; cholesterol metabolism;  
XX steroid; post-translational modification; gene therapy;  
XX rhabdomyolysis; cell morphology; neoplasia; cancer.  
XX Homo sapiens.  
XX OS  
XX PN

AWM22498	ID	AAW22498	standard; Protein; 376 AA.
XX	AC	AAW22498;	
XX	DE	10-MAR-1998	(first entry)
XX	DE	Phaffia derived carotenoid biosynthesis pathway enzyme PrcrE.	
XX	KW	Glyceraldehyde-3-phosphate dehydrogenase; GAPDH; carotenoid;	
XX	KW	synthesis; promoter; recombinant DNA; astaxanthin; ribosomal protein;	
XX	KW	food colouring.	
XX	OS	Phaffia rhodozyma.	
XX	PN	MO9723633-A1.	
XX	PD	03-JUL-1997.	
XX	PF	23-DEC-1996; 96WO-EP05887.	
XX	PR	11-APR-1996; 96EP-0200943.	
XX	PR	22-DEC-1995; 95EP-0203620.	
XX	PA	(KONN ) GIST-BROCADES BV.	
XX	PI	(COIJ/) COIJEN A J J.	
XX	PI	Verdoes JC, Wery J;	
XX	DR	WPI: 1997-351068/32.	
XX	DR	N-PSDB; AAT72942.	
XX	PT	Phaffia derived GAPDH and carotenoid synthesis genes and promoter	
XX	PT	fragment - used in the recombinant production of therapeutically	
XX	PT	useful proteins e.g. carotenoids for use in food colouring	
PS	Claim 11;	Page 55-56; 118pp; English.	
XX	CC	The present sequence represents a Phaffia derived carotenoid	
XX	CC	biosynthesis pathway enzyme. The nucleic acid encoding this protein	
XX	CC	can be used in the novel recombinant DNA of the present invention. The	
XX	CC	recombinant DNA comprises a transcription promoter operably linked to a	
XX	CC	downstream sequence to be expressed, where the transcription promoter	
XX	CC	comprises a region found upstream of the open reading frame (ORF) of a	
XX	CC	highly expressed Phaffia gene (preferably GAPDH, a ribosomal protein	
XX	CC	gene, an enzyme involved in the biosynthesis pathway). The recombinant	
XX	CC	DNA can be used to transform hosts, preferably Phaffia. These	
XX	CC	transformed hosts are then used in the recombinant production of GAPDH	
XX	CC	or an enzyme involved in carotenoid synthesis, preferably astaxanthin.	
XX	CC	They may also be used to produce a pharmaceutical product. Purified	
XX	CC	carotenoids can be used as colourants in food and/or feed, and also in	
XX	CC	cosmetics.	
XX	Sequence	376 AA;	
50	Query Match	45.1%; Score 707; DB 18; Length 376;	
	Best Local Similarity	41.9%; Pred. No. 5.5e-60;	
	Matches 148; Conservative	57; Mismatches 76; Indels 72; Gaps 5;	
QY	11	ILLEFYKYLLOLPGKQVTRTKLSQAFNHMLKVPEDKIKQIIIEVTENMLHNASLIDIEDNS 70	
Db	21	VLLPEYHYLgkmpgkretisqgleatnhyldvkkedlevlgnvgnhlhtcasllmddveds 80	
QY	71	KIRKGFPVAHSHYTGIPSYNSANYVYFELGKEKVLTLDBPDV----- 112	
Db	81	vlrtspvahllylpgtlnntanyyflaygelfkl-rptlpmpvlpssaslgssvss 139	
QY	113	-----KLFTRQLELHGOGGLDIYWRNRYCPTFE 142	
Db	140	assssassengytsltnsqipfskdyldkvtldemslshrggglfelftwrdsltcpsce 199	
QY	143	EYKANVLKTKGTLGLANGLMOLFSDYKEDDKPLNLNTGLFQIQRDDYANLHRSKEYSNK 202	



Db 200 eykvnvlgktgllfrfavrlnmaksecidifvqlvnlislyfqrldymnlgssayahnk 259  
Oy 203 SPCEDLTGKSPFPTIHMWSRPESTOVONILKRTENIDIKKYVHLE-DVGSFETR 261  
Db 260 nfaeltsgktsifpctihnpssrlyvnlqktspeilhcnymrcthsfeytq 319  
Oy 262 NLKLELE-----AKAYKQID-----ARGGNPELVLVKHL 292  
Db 320 evlnlslgalereigrigfgefaansrmdlgdvdsgrtgrkvnkllaikla 372

## RESULT 7

AAW12391

ID AAW12391 standard; Protein: 330 AA.

AC AAW12391;

DT 17-JUN-1997 (first entry)

DE Geranylgeranyl diphosphate synthase F77S/Y101H mutant.

Geranylgeranyl diphosphate synthase; GGPS; mutagenesis;  
enzyme engineering; protein engineering; prenyl diphosphate;  
steroid; carotenoid; vitamin E; vitamin K; ubiquinone;  
Sulfolobus acidocaldarius; geranylflarnesyl diphosphate.

OS Synthetic.

PN EP763542-A2.

PD 19-MAR-1997.

PE 30-AUG-1996; 96EP-0113930.

PF 01-SEP-1995; 95JP-0247043.

PR (TOYT ) TOYOTA JIDOSHA KK.

PA Asada C, Hemmi H, Hirooka K, Nishino T, Ohnuma S;

PI Ohto C;

DR WPI: 1997-167753/16.

N-PSDB; AAT63242.

PT New mutant geranylgeranyl diphosphate synthase enzymes - are

capable of forming prenyl diphosphate with a longer chain length

than that produced by the native enzyme

PS Disclosure; Page 21-23; 29pp; English.

A mutated geranylgeranyl diphosphate synthase (GGPS) (AAW12391) has  
Phe77Ser and Tyr101His amino acid substitutions in relation to the  
CC native GGPS (AAW12386) of Sulfolobus acidocaldarius. It is encoded  
CC by a mutant gene (AAT63242) obtd. by random mutagenesis of the native  
CC GGPS gene (AAT63237) using nitrite. 5 Mutant GGPS enzymes (AAW12387-  
CC 911) have been produced, that are capable of forming prenyl  
CC diphosphates having longer chain lengths than those produced by the  
CC native enzyme, e.g. geranylflarnesyl diphosphate (C25).

XX Sequence 330 AA;

Query Match 13.5%; Score 212; DB 18; Length 330;

Best Local Similarity 24.7%; Pred. No. 2.8e-12;

Matches 74; Conservative 55; Mismatches 128; Indels 42; Gaps 10;

Oy 12 LLEPKYLLQLGKQVRRKLSQAFNHMLKVPEDKQIIEVEMLNHNSLLDIEDNSK 71  
Db 30 lyeasyhltsggkrlrpllltlssdlfggqreayagaalevlhtstclvndldmqdn 89  
Oy 72 LRRGFPVAHSIGIPSVINSANYV---FLGLEKVLTDHPDAVK---LFTROLLELHQ 124

Db 90 lrrgipvtvkhgjlplallegqllhakaftqlltgalrglpsetlikafdtfrsllise 149  
Oy 125 GGGIDYWRDNTCTPEEYKAMVQKTKGLEGGLAVGL-----MQLFSYKRD 173  
Db 150 gqavmefedridt-kegeyldmistrktgaltsasssfgallaganddvrlmsdfgrn- 207  
Oy 174 KPLNLTGLFQIRDDYANLHSEKSEKSPCEDTTEGKFSPTIHA--IMSREPTQVQ 231  
Db 208 -----lgaifqvddllgltadekekgkpyfsdliregkktllvktlelckedekkiyl 261  
Oy 232 NLKRTENID-----IKKYCVHLEDVGSFEYTRNTLKELEAKAYKQIDARGNP 282  
Db 262 kalgntksaskelmsadlikkysldygnlae-kyxnaldsl-----ngysksdip 314

## RESULT 8

AAW12389

ID AAW12389 standard; Protein: 330 AA.

AC AAW12389;

DT 17-JUN-1997 (first entry)

DE Geranylgeranyl diphosphate synthase F77S mutant.

Geranylgeranyl diphosphate synthase; GGPS; mutagenesis;  
enzyme engineering; protein engineering; prenyl diphosphate;  
steroid; carotenoid; vitamin E; vitamin K; ubiquinone;  
Sulfolobus acidocaldarius; geranylflarnesyl diphosphate.

OS Synthetic.

PN EP763542-A2.

PD 19-MAR-1997.

PE 30-AUG-1996; 96EP-0113930.

PF 01-SEP-1995; 95JP-0247043.

PR (TOYT ) TOYOTA JIDOSHA KK.

PA Asada C, Hemmi H, Hirooka K, Nishino T, Ohnuma S;

PI Ohto C;

DR WPI: 1997-167753/16.

N-PSDB; AAT63240.

PT New mutant geranylgeranyl diphosphate synthase enzymes - are

capable of forming prenyl diphosphate with a longer chain length

than that produced by the native enzyme

PS Disclosure; Page 17-19; 29pp; English.

A mutated geranylgeranyl diphosphate synthase (GGPS) (AAW12389) has  
a Phe77Ser substn. in relation to the native GGPS (AAW12386) of  
CC Sulfolobus acidocaldarius. It is encoded by a mutant gene (AAT63240)  
CC obtd. by random mutagenesis of the native GGPS gene (AAT63237). 5  
CC Mutant GGPS enzymes (AAW12387-91) have been produced that are capable  
CC of forming prenyl diphosphates having longer chain lengths than  
CC those produced by the native enzyme, e.g. geranylflarnesyl  
CC diphosphate (C25).

XX Sequence 330 AA;

Query Match 13.5%; Score 211; DB 18; Length 330;

Best Local Similarity 24.5%; Pred. No. 3.4e-12;

Matches 72; Conservative 52; Mismatches 124; Indels 46; Gaps 9;

Oy 12 LLEPKYLLQLGKQVRRKLSQAFNHMLKVPEDKQIIEVEMLNHNSLLDIEDNSK 71  
Db 30 lyeasyhltsggkrlrpllltlssdlfggqreayagaalevlhtstclvndldmqdn 89

QY 72 LRGEPAVASHYIGIPSVINSANVYFLGLEKVLTDHP-----DAVKLFTROLLE 121  
 Db 90 IRRGLPVRHVKYGLPALLAGDLHAKAFGL---LITGLRGLPSETLIKAFDIFTRSI 146  
 QY 122 LHOOGGLDIYWRDNTCTEETEEKYAMVLOKGTGFLGVLG-----MOLFSDYK 170  
 Db 147 ISEGVAVMEFEDRID-KEGEYIDMISRTKTAALFSASSIGALLAGANDDVRLMSDFG 205  
 QY 171 EDLPLMTLGLFQIRDDYANLHSEKSEKSPCEDLTEGKFSFPTIHA--IWSRPEST 228  
 Db 206 -----TNIATGIVDDLLGLTADKEKELGKPVSDIREGKKTLLIVIKLELCKEDKK 258  
 QY 229 QVONILRQRTENID-----IKKYCVHYLEDVGSFEYTRN---TKELPAKA 271  
 Db 259 IVLKALGNKSAKSEELMSADILKKYSLDYAYNLAE-KYKNALDSLNGLSKSS 311

## RESULT 9

AAW12390  
 ID AAW12390 standard; Protein; 330 AA.

AC AAW12390;

DT 17-JUN-1997 (first entry)

DE Geranylgeranyl diphosphate synthase F77L/V99M mutant.

Geranylgeranyl diphosphate synthase; GGPS; mutagenesis;  
 enzyme engineering; protein engineering; prenyl diphosphate;  
 steroid; carotenoid; vitamin E; vitamin K; ubiquinone;  
 Sulfolobus acidocaldarius; geranylarnesyl diphosphate.

OS Synthetic.

PN EP63542-A2.

PD 19-MAR-1997.

PE 30-AUG-1996; 96EP-0113930.

PR 01-SEP-1995; 95JP-0247043.

PA (TOYT ) TOYOTA JIDOSHA KK.

PI Asada C, Hemmi H, Hirooka K, Nishino T, Ohnuma S;

PI Ohto C;

DR WPI: 1997-167753/16.

DR N-PSDB; AAT63241.

PT New mutant geranylgeranyl diphosphate synthase enzymes - are

capable of forming prenyl diphosphate with a longer chain length

than that produced by the native enzyme

Disclosure; Page 19-21; 29pp; English.

PS A mutated geranylgeranyl diphosphate synthase (GGPS) (AAW12390) has  
 CC Phe77Leu and Val199Met amino acid substitutions in relation to the  
 CC native GGPS (AAW12386) of Sulfolobus acidocaldarius. It is encoded  
 CC by a mutant gene (AAT63241) obtd. by random mutagenesis of the native  
 CC GGPS gene (AAT63237) using nitrite. 5 Mutant GGPS enzymes (AAW12387-  
 CC 91) have been produced that are capable of forming prenyl  
 CC diphosphates having longer chain lengths than those produced by the  
 CC native enzyme, e.g. geranylarnesyl diphosphate (C25).

XX Sequence 330 AA;

Query Match 13.3%; Score 209; DB 18; Length 330;  
 Best Local Similarity 24.7%; Pred. No. 5,4e-12;  
 Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

QY 12 LLEPKYLLDLPCKQVTRKLSQAFNHWLKVPEDEKIQIIIEVTEMLHNASLLIDIEDNSK 71  
 Db 30 Lyeasylfsggkrllpillsdlfggqreyagaievlhcltlvnddmqdn 89  
 QY 72 LRGEPAVASHYIGIPSVINSANVY---FLGLEKVLN-LDHPDAVK---LFTROLLEHQ 124  
 Db 90 IRRGLPVRHVKYGLPALLAGDLHAKAFGLLQALRGLPSETLIKAFDIFTRSI 149  
 QY 125 GGGGLDIYWRDNTCTEETEEKYAMVLOKGTGFLGVLG-----MOLFSDYKEDL 173  
 Db 150 GGVAVMEFEDRID-KEGEYIDMISRTKTAALFSASSIGALLAGANDDVRLMSDFG 205  
 QY 174 KPLINTLGLFQIRDDYANLHSEKSEKSPCEDLTEGKFSFPTIHA--IWSRPESTQVQ 231  
 Db 206 -----TNIATGIVDDLLGLTADKEKELGKPVSDIREGKKTLLIVIKLELCKEDKK 261  
 QY 232 NILRQRTENID-----IKKYCVHYLEDVGSFEYTRNLTKELEAKAYQIDARGGNP 282  
 Db 262 KALGNKSAKSEELMSADILKKYSLDYAYNLAE-KYKNALDSI-----QVSKSDIP 314

## RESULT 10

AAW12386  
 ID AAW12386 standard; Protein; 330 AA.

AC AAW12386;

DT 17-JUN-1997 (first entry)

DE Geranylgeranyl diphosphate synthase.

Geranylgeranyl diphosphate synthase; GGPS; mutagenesis;  
 enzyme engineering; protein engineering; prenyl diphosphate;  
 steroid; carotenoid; vitamin E; vitamin K; ubiquinone.

OS Sulfolobus acidocaldarius (ATCC-33909).

PN EP63542-A2.

PD 19-MAR-1997.

PE 30-AUG-1996; 96EP-0113930.

PR 01-SEP-1995; 95JP-0247043.

PA (TOYT ) TOYOTA JIDOSHA KK.

PI Asada C, Hemmi H, Hirooka K, Nishino T, Ohnuma S;

PI Ohto C;

DR WPI: 1997-167753/16.

DR N-PSDB; AAT63237.

PT New mutant geranylgeranyl diphosphate synthase enzymes - are

capable of forming prenyl diphosphate with a longer chain length

than that produced by the native enzyme

Disclosure; Page 11-13; 29pp; English.

PS The wild-type geranylgeranyl diphosphate synthase (GGPS) (AAW12386)  
 CC of Sulfolobus acidocaldarius is able to synthesize prenyl  
 CC diphosphates that function as precursors of steroids, carotenoids,  
 CC as transition substrates of prenylated proteins, and as substrates  
 CC for the synthesis of vitamin E. Vitamin K, ubiquinone, etc. Mutant  
 CC GGPS enzymes have been produced (see also AAW12387-91) in which  
 CC Phe77, Met85, Val99, Tyr101, Phe118, Arg119 and/or Asp312 are  
 CC substd. by another amino acid. These mutants are capable of  
 CC forming prenyl diphosphates of at least 25 C atoms, such as  
 CC geranylarnesyl diphosphate; the native enzyme is unable to form  
 CC such long-chain prenyl diphosphates.

XX Sequence 330 AA;

Query Match 13.3%, Score 208; DB 18; Length 330;  
 Best Local Similarity 24.7%, Pred. No. 6,76-12;  
 Matches 74; Conservative 53; Mismatches 130; Indels 42; Gaps 10;

OY 12 LLEPKYLLQLPKGVKRLKSOAFNHLKVPEDKQIIIEVTEMLHNSLIDIEDNSK 71  
 Db 30 lyeaayhlftsgkrlrlpdlitissdlfggqreayagaalevhtftlyvndmdqdn 89  
 OY 72 LRRGPRVASHSYGIPSVNSANYV---FLGLEKVLTDHPDAK---LFTROLLEHQ 124  
 Db 90 lrrgprlvhvkypdlallagdlhakaftqlltgalrlpsetllkardiftrslilise 149  
 OY 125 GGGLDIWRDNTCTPEEYKAMVLQKTGGLFGLAVGL-----MQLFSDYKEDL 173  
 Db 150 gqavmefedridt-kegeyldmisktaalfsaassigallagandvrlmsdfg--- 205  
 OY 174 KPLNTLGLFQIRDVANLHSEKSEKSFCEDLTEGKFSFPTIHA--IMSPESTOVQ 231  
 Db 206 ---enlglafqvddllgltdakekqlgkpvfsdiregkklvllktlelckedeckivl 261  
 OY 232 NILKROTENID-----IKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNP 282  
 Db 262 kalgnksaskeelmsadlkkysldayunlae-kyknaldsl-----ngvssksdip 314

## RESULT 11

AAR79969 standard; Protein; 330 AA.

AAR79969;

25-APR-1996 (first entry)

Geranylgeranyl diphosphate synthase.

Geranylgeranyl diphosphate synthase; GGP; thermostable; isoprenoid;  
 isopentenyl diphosphate; farnesyl diphosphate; caloteneid; diisoprene;  
 vitamin; pmalcgl; Sulfolobus acidocaldarius.

Sulfolobus acidocaldarius.

Ep674000-A2.

27-SEP-1995.

23-MAR-1995; 95EP-0104301.

25-NOV-1994; 94JP-0315572.

24-MAR-1994; 94JP-0053804.

(TOYT ) TOYOTA JIDOSHA KK.

Asada C, Higuchi Y, Nishino T, Ohnuma S, Ohto C;

Suzuki M, Takeuchi Y;

WPI: 1995-32867/43.

N-PSDB; AAT04237.

DNA encoding heat stable geranylgeranyl di:phosphate synthase  
 (GGP) from Sulfolobus acidocaldarius - useful for synthesis of  
 GGP, an intermediate for carotenoid production.

Claim 2; Page 11-13; 21pp; English.

This sequence represents thermostable geranylgeranyl diphosphate (GGP)  
 synthase. GGP is synthesised in vivo by condensation of isopentenyl  
 diphosphate and farnesyl diphosphate, and is an important intermediate  
 for biosynthesis of isoprenoids and isoprenoid containing compounds. The  
 DNA encoding this sequence was inserted into a plasmid vector (termed  
 pmalcgl) which was used to transform E.coli cells to produce the GGP  
 synthase. This GGP synthase is more stable towards heat than similar  
 enzymes from mesophilic organisms. By using this plasmid, GGP synthase

CC can therefore be produced at low cost.  
 XX Sequence 330 AA;  
 SQ

Query Match 13.2%, Score 207; DB 16; Length 330;  
 Best Local Similarity 24.7%, Pred. No. 8,4e-12;  
 Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

OY 12 LLEPKYLLQLPKGVKRLKSOAFNHLKVPEDKQIIIEVTEMLHNSLIDIEDNSK 71  
 Db 30 lyeaayhlftsgkrlrlpdlitissdlfggqreayagaalevhtftlyvndmdqdn 89  
 OY 72 LRRGPRVASHSYGIPSVNSANYV---FLGLEKVLTDHPDAK---LFTROLLEHQ 124  
 Db 90 lrrgprlvhvkypdlallagdlhakaftqlltgalrlpsetllkardiftrslilise 149  
 OY 125 GGGLDIWRDNTCTPEEYKAMVLQKTGGLFGLAVGL-----MQLFSDYKEDL 173  
 Db 150 gqavmefedridt-kegeyldmisktaalfsaassigallagandvrlmsdfg--- 205  
 OY 174 KPLNTLGLFQIRDVANLHSEKSEKSFCEDLTEGKFSFPTIHA--IMSPESTOVQ 231  
 Db 206 ---enlglafqvddllgltdakekqlgkpvfsdiregkklvllktlelckedeckivl 261  
 OY 232 NILKROTENID-----IKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNP 282  
 Db 262 kalgnksaskeelmsadlkkysldayunlae-kyknaldsl-----ngvssksdip 314

## RESULT 12

AAM12387 standard; Protein; 330 AA.

AAM12387;

17-JUN-1997 (first entry)

Geranylgeranyl diphosphate synthase M851/R199K/D312N mutant.

Geranylgeranyl diphosphate synthase; GGPS; mutagenesis;  
 enzyme engineering; protein engineering; prenyl diphosphate;  
 steroid; carotenoid; vitamin E; vitamin K; ubiquinone;  
 Sulfolobus acidocaldarius; geranyl farnesyl diphosphate.

Synthetic.

Ep763542-A2.

19-MAR-1997.

30-AUG-1996; 96EP-0113930.

01-SEP-1995; 95JP-0247043.

(TOYT ) TOYOTA JIDOSHA KK.

Asada C, Hemmi H, Hirooka K, Nishino T, Ohnuma S;

Ohto C;

WPI: 1997-167753/16.

N-PSDB; AAT63238.

New mutant geranylgeranyl diphosphate synthase enzymes are  
 capable of forming prenyl diphosphate with a longer chain length  
 than that produced by the native enzyme

Disclosure: Page 13-15; 29pp; English.

A mutated geranylgeranyl diphosphate synthase (GGPS) (AAM12387) has  
 M851, R199K and D312N amino acid substitutions in relation to the  
 native GGPS (AAM12386) of Sulfolobus acidocaldarius. It is encoded  
 by a mutant gene (AAT63238) obtd. by random mutagenesis of the native

PS Disclosure; Pages 29-30; 47pp; English.

```

0Y      12 LLEPRKYLLOLPGKOVFRKLDSOAFENHMYLKVEBDEKQIILEVTEWELAHNNAISLLIDIEONSK 71
        | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     30 lyeasynhlftsgykrllrllltissdlfggqreayagaiaeivlnhtflvhdmdmqn 89
0Y      72 LRERGEVAHSIYGIDPSVINSANYV---FLGLEKYLT-LDHPDAVK---LETROLLEHQ 124
        | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     90 irrgptvnyvkgyrlallaaqdlhnakafiglltgaqlgsetlilkafidfrzllisse 149
0Y      125 GGGIDIVMNDNTCTEERBYKAMVLORKGIFGLAVG-----MOLFSDYKEDL 173
        | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     150 gaaavameledridt-kegeyalmdsrkkaalffaassigallaagandndvrlmedfg--- 205
0Y      174 KPLNLTLGLFQIRDDVYANLSKEYSENKSFCEDLTEGRFSPTTIA--IWSRPESTVOQ 231
        | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     206 ----tnlgatfvddllgladekelgkpyfsdrrgekkctlllyktlelokedekkiyl 261
0Y      232 NLLRRTENID-----IKKCYCHLEDVGSEFYTNNTLKELEAKAYKOIIDRGNGP 282
        | | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     262 kalgnksakeelmssadikkysldygmnae-kyyrnaadsl----nqvsskatdp 314

```

PS Claim 8; Page 13-14; 24pp; English

XX This sequence is the Sulfolobus acidocaldarius geranylgeranyl  
 CC diphosphate synthase. This sequence is mutated to give the mutant prenyl  
 CC diphosphate synthase of the invention. The mutant enzyme has a modified  
 CC aspartic acid-rich domain in region II. The enzyme is useful for  
 CC producing a prenyl diphosphate having not more than 15 carbons,  
 CC especially farne syl diphosphate, by contacting the enzyme with a  
 CC substrate selected from isopentenyl diphosphate, dimethylallyl  
 CC diphosphate and geranyl diphosphate. Prenyl diphosphates are precursors  
 CC of steroids, ubiquinones, dolichols, carotenoids, prenylated proteins,  
 CC animal hormones, plant hormones, etc. The mutations impart a higher  
 CC thermal stability to the enzyme and a higher specificity.

CC Sequence 330 AA;

Query Match 13.2%; Score 207; DB 19; Length 330;  
 Best Local Similarity 24.7%; Pred. No. 8.4e-12;  
 Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

OY 12 LLEPKYLLQLPGKQVFRKLSQAFNHMLKVPEDKLIQIIIEVTMLHNSLIDIEDNSK 71  
 DB 30 Iyeasyhlftsggkrlrlpilitissdlfggrerayagaalevlhtftlvhddmdqdn 89  
 OY 72 LRRGFPVAHSITGIPSVINSANYVY---FLGLEKVLV-LDHPDAVK---LFTROLLEHQ 124  
 DB 90 Irtgipvtvkvgyrlpallagallnakafqlltgalrtgipetlika dftlrslilise 149  
 OY 125 GGGLDIYWRDNTCTPEEYKAMVLQKTGGLGLAVGL-----MOLFSDYKEDL 173  
 DB 150 ggavmefedridl-kegeyldmlsrktaalfsasssigaallaganddvrlmsdfg--- 205  
 OY 174 KPLMTLGLFQIRDDYANLSKEYSKESKCEDLTEGKFSPTTHA--IMSREPTOVQ 231  
 DB 206 ----tnlialafivddilglitadekelgkpvfsdrtregkklvlylktlelckedekklvl 261  
 OY 232 NILRORTENID-----IKRYCVHYLEDVGSFEYTRNTLKELEPAKAYKQIDARGNP 282  
 DB 262 kalgnskskeelmsaadlikkysldaynlae-kyyknalnsdl-----ngvssksdip 314

RESULT 15

AAM52283 ID AAM52283 standard; Protein; 330 AA.

XX AAM52283;

XX 22-JUN-1998 (first entry)

XX Mutant prenyl diphosphate synthase.

XX Geranylgeranyl diphosphate synthase; prenyl diphosphate synthase; enzyme;

KW aspartic acid-rich domain; precursor: steroid; ubiquinone; dolichol;

XX carotenoid; prenylated protein; hormone.

OS Sulfolobus acidocaldarius.

XX Key location/qualifiers

FT Misc-difference 77 /label- F77Y

FT Misc-difference 78 /note- "Phe to Tyr mutation"

FT Misc-difference 80 /label- T78X

FT Misc-difference 81 /note- "X- Phe, or Ser"

FT Misc-difference 81 /label- V80F

FT Misc-difference 84 /note- "Val to Phe mutation"

FT Misc-difference 84 /label- H81X

FT Misc-difference 84 /note- "X- Leu, or Ala"

FT Misc-difference 84 /label- 184L

FT /note- "Ile to Leu mutation"

XX EP821065-A2.  
 PN 28-JAN-1998.  
 PD 16-JUL-1997; 97EP-0112144.  
 PF 24-JUL-1996; 96JP-0213211.  
 PR (TOYT ) TOYOTA JIDOSHA KK.  
 PA Hirooka K, Nakane H, Nishino T, Ohnuma S, Ohto C;  
 PI WPI; 1998-102621/10.

PT Mutant prenyl diphosphate synthase - especially useful for  
 PT producing farne syl diphosphate  
 PS Claim 9; Page -: 24pp; English.

CC This sequence is a mutant prenyl diphosphate synthase of the invention,  
 CC and was produced by mutating the Sulfolobus acidocaldarius geranylgeranyl  
 CC diphosphate synthase. The mutant enzyme has a modified aspartic acid-rich  
 CC domain in region II. The enzyme is useful for producing a prenyl  
 CC diphosphate having not more than 15 carbons, especially farne syl  
 CC diphosphate, by contacting the enzyme with a substrate selected from  
 CC isopentenyl diphosphate, dimethylallyl diphosphate and geranyl  
 CC diphosphate. Prenyl diphosphates are precursors of steroids, ubiquinones,  
 CC dolichols, carotenoids, prenylated proteins, animal hormones, plant  
 CC hormones, etc. The mutations impart a higher thermal stability to the  
 CC enzyme and a higher specificity.

XX Sequence 330 AA;

Query Match 13.2%; Score 207; DB 19; Length 330;  
 Best Local Similarity 24.4%; Pred. No. 8.4e-12;  
 Matches 73; Conservative 52; Mismatches 132; Indels 42; Gaps 10;

OY 12 LLEPKYLLQLPGKQVFRKLSQAFNHMLKVPEDKLIQIIIEVTMLHNSLIDIEDNSK 71  
 DB 30 Iyeasyhlftsggkrlrlpilitissdlfggrerayagaalevlhtftlvhddmdqdn 89  
 OY 72 LRRGFPVAHSITGIPSVINSANYVY---FLGLEKVLV-LDHPDAVK---LFTROLLEHQ 124  
 DB 90 Irtgipvtvkvgyrlpallagallnakafqlltgalrtgipetlika dftlrslilise 149  
 OY 125 GGGLDIYWRDNTCTPEEYKAMVLQKTGGLGLAVGL-----MOLFSDYKEDL 173  
 DB 150 ggavmefedridl-kegeyldmlsrktaalfsasssigaallaganddvrlmsdfg--- 205  
 OY 174 KPLMTLGLFQIRDDYANLSKEYSKESKCEDLTEGKFSPTTHA--IMSREPTOVQ 231  
 DB 206 ----tnlialafivddilglitadekelgkpvfsdrtregkklvlylktlelckedekklvl 261  
 OY 232 NILRORTENID-----IKRYCVHYLEDVGSFEYTRNTLKELEPAKAYKQIDARGNP 282  
 DB 262 kalgnskskeelmsaadlikkysldaynlae-kyyknalnsdl-----ngvssksdip 314

Search completed: August 19, 2001, 11:22:34  
 Job time: 118 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 19, 2001, 11:21:21 ; Search time 12.27 Seconds  
(without alignments)  
503,431 Million cell updates/sec

Title: US-09-744-527-4

Perfect score: 300  
Sequence: 1 MEKQETVQRILPEYKYL.....NPELVAVKHLKMKKEENE 300

Scoring table: OLIGO  
Gapop 60.0 , Capext 60.0

Searched: 197339 seqs, 20590346 residues

Word size : 0

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2.6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgn2.6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgn2.6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgn2.6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgn2.6/ptodata/2/iaa/PCPUS.COMB.pep:\*  
6: /cgn2.6/ptodata/2/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	203	67.7	300	1	US-08-469-665-2
2	203	67.7	300	2	US-09-038-596-2
3	203	67.7	300	5	PCT-US95-00421-2
4	10	3.3	335	2	US-08-761-344-2
5	7	2.3	119	2	US-08-710-749-27
6	7	2.3	368	2	US-08-903-624-3
7	7	2.3	368	4	US-08-973-914-9
8	7	2.3	699	4	US-08-851-843A-52
9	7	2.3	699	4	US-08-974-549A-188
10	7	2.3	719	4	US-08-851-843A-17
11	7	2.3	719	4	US-08-974-549A-219
12	6	2.0	6	1	US-08-347-792-8
13	6	2.0	6	1	US-08-431-357-8
14	6	2.0	6	3	US-08-697-221-9
15	6	2.0	6	3	US-08-920-610-6
16	6	2.0	6	5	PCT-US95-15353-8
17	6	2.0	9	4	US-09-217-609A-11
18	6	2.0	9	4	US-08-873-235B-11
19	6	2.0	23	3	US-09-120-365-36
20	6	2.0	23	4	US-09-515-039-36
21	6	2.0	45	2	US-08-245-511-56
22	6	2.0	45	2	US-08-600-993A-56
23	6	2.0	48	4	US-09-177-249-202
24	6	2.0	60	4	US-08-928-213B-68
25	6	2.0	60	4	US-08-928-213B-72
26	6	2.0	82	1	US-08-497-312-30
27	6	2.0	87	4	US-09-284-033-3

28	6	2.0	87	4	US-08-729-834B-3	Sequence 3, Appl
29	6	2.0	96	1	US-08-202-389-33	Sequence 33, Appl
30	6	2.0	122	6	5169835-12	Patent No. 5169835
31	6	2.0	143	6	5169835-18	Patent No. 5169835
32	6	2.0	144	6	5169835-8	Patent No. 5169835
33	6	2.0	160	1	US-08-167-035-40	Sequence 40, Appl
34	6	2.0	160	1	US-08-208-887A-40	Sequence 40, Appl
35	6	2.0	160	2	US-08-539-005-40	Sequence 40, Appl
36	6	2.0	173	4	US-09-133-341-13	Sequence 14, Appl
37	6	2.0	176	4	US-09-133-341-14	Sequence 14, Appl
38	6	2.0	178	3	US-08-637-759B-225	Sequence 225, App
39	6	2.0	178	3	US-08-871-355A-225	Sequence 225, App
40	6	2.0	207	4	US-09-068-960-43	Sequence 43, Appl
41	6	2.0	210	1	US-08-730-895A-2	Sequence 2, Appl
42	6	2.0	210	4	US-09-099-613-2	Sequence 2, Appl
43	6	2.0	220	5	PCT-US95-03236-36	Sequence 3, Appl
44	6	2.0	248	6	5169835-15	Patent No. 5169835
45	6	2.0	270	2	US-08-859-201-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-469-665-2  
Sequence 2, Application US/08469665  
Patent No. 5786193  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,665  
FILING DATE: June 6, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: 11 JAN 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-377  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS: LINEAR  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-469-665-2

Query Match 67.7%; Score 203; DB 1; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.5e-186;  
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 MEKQETVQRILPEYKYLQLPCKQVYRTKLSQAFNHLKVPEDKLIITIEVTMLNNS 60

Db 1 MEKTOETVORILLEPYKYLLOLPKQVTRKLSQAFNHMLKVPEDKLOIIIEVTMLHNAS 60  
QY 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
Db 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
QY 181 GLFQIRDDYANLHSEKSEKNS 203  
Db 181 GLFQIRDDYANLHSEKSEKNS 203

RESULT 2  
US-09-596-2  
Sequence 2, Application US/09038596  
Patent No. 5928924  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/038, 596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/469, 665  
FILING DATE: June 6, 1995  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: 11 JAN 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-377  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-09-038-596-2

Query Match 67.7%; Score 203; DB 2; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.5e-186; Indels 0; Gaps 0;  
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKTOETVORILLEPYKYLLOLPKQVTRKLSQAFNHMLKVPEDKLOIIIEVTMLHNAS 60  
Db 1 MEKTOETVORILLEPYKYLLOLPKQVTRKLSQAFNHMLKVPEDKLOIIIEVTMLHNAS 60  
QY 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
Db 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
QY 181 GLFQIRDDYANLHSEKSEKNS 203  
Db 181 GLFQIRDDYANLHSEKSEKNS 203

Db 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
QY 181 GLFQIRDDYANLHSEKSEKNS 203  
Db 181 GLFQIRDDYANLHSEKSEKNS 203

RESULT 3  
PCT-US95-00421-2  
Sequence 2, Application PC/TUS9500421  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: Concurrently  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-257  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
PCT-US95-00421-2

Query Match 67.7%; Score 203; DB 5; Length 300;  
Best Local Similarity 100.0%; Pred. No. 1.5e-186; Indels 0; Gaps 0;  
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEKTOETVORILLEPYKYLLOLPKQVTRKLSQAFNHMLKVPEDKLOIIIEVTMLHNAS 60  
Db 1 MEKTOETVORILLEPYKYLLOLPKQVTRKLSQAFNHMLKVPEDKLOIIIEVTMLHNAS 60  
QY 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
Db 61 LLIIDIEDNSKLRGEPVAHSIYGIPSVINSANYVFLGLEKVLTLDPDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTTEEEKAMVLOKTGGLFGLAVGLMQLFSDYKEDLKPILNTL 180  
QY 181 GLFQIRDDYANLHSEKSEKNS 203  
Db 181 GLFQIRDDYANLHSEKSEKNS 203



Db 181 GLEKRDYANLHRSKEYSENKS 203

RESULT 4  
US-08-761-344-2  
Sequence 2, Application US/08761344

Patent No. 5912154

GENERAL INFORMATION:

APPLICANT: Ferro-No. 5912154ick, Susan

APPLICANT: Jiang, Yu

TITLE OF INVENTION: GERANYLGERANYL DIPHOSPHATE SYNTHASE

TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street

CITY: Denver

STATE: CO

COUNTRY: U.S.A.

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/761,344

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 30,020

REFERENCE/DOCKET NUMBER: 3161-14

TELEPHONE: 303/863-9700

TELEFAX: 303/862-0223

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 335 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-761-344-2

Query Match 3.3%; Score 10; DB 2; Length 335;

Best Local Similarity 100.0%; Pred. No. 0.081;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 SLIDDIEDN 69

Db 71 SLIDDIEDN 80

RESULT 5

US-08-710-749-27

Sequence 27, Application US/08710749

Patent No. 5955089

GENERAL INFORMATION:

APPLICANT: Briles, David E.

APPLICANT: Hollingshead, Susan

TITLE OF INVENTION: STRAIN SELECTION OF PNEUMOCOCCAL SURFACE

TITLE OF INVENTION: PROTEINS

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Curtis, Morris & Safford

STREET: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/710,749

FILING DATE: 20-SEP-1996

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2074

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

TELEFAX: (212) 840-0712

INFORMATION FOR SEQ ID NO: 27:

SEQUENCE CHARACTERISTICS:

LENGTH: 119 amino acids

TYPE: amino acid

STRANDEDNESS: n/a

TOPOLOGY: linear

MOLECULE TYPE: amino acid

US-08-710-749-27

Query Match 2.3%; Score 7; DB 2; Length 119;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 LGLEKVL 104

Db 6 LGLEKVL 12

RESULT 6

US-08-903-624-3

Sequence 3, Application US/08903624

Patent No. 5877013

GENERAL INFORMATION:

APPLICANT: Liao, Gwo-Jen

APPLICANT: Lee, Yi-Jang

APPLICANT: Lee, Yun-Huey

APPLICANT: Chen, Li-Lin

APPLICANT: Chu, Wen-Shen

TITLE OF INVENTION: RHODOSPORIDIUM D-AMINO ACID OXIDASE

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Fish & Richardson, P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/903,624

FILING DATE: July 30, 1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:

NAME: Tsao, Y. Rocky

REGISTRATION NUMBER: 34,053

REFERENCE/DOCKET NUMBER: 06497/006001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 368 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-903-624-3

Query Match  
Best Local Similarity 2.3%; Score 7; DB 2; Length 368;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ETVORIL 12  
DB 250 ETVORIL 256

RESULT 7  
US-08-973-914-9  
Sequence 9, Application US/08973914F  
Patent No. 6187574  
GENERAL INFORMATION:  
APPLICANT: GARCIA LOPEZ, Jose Luis  
APPLICANT: CORTES RUBIO, Estrella  
APPLICANT: BARREDO FUENTE, Jose Luis  
APPLICANT: DIEZ GARCIA, Bruno  
APPLICANT: MORENO VALLE, Miguel Angel  
APPLICANT: SCHLEISSNER SANCHEZ, Carmen  
APPLICANT: COLLADOS DE LA VIEJA, Alfonso  
APPLICANT: SALTO MALDONADO, Francisco  
TITLE OF INVENTION: PROCESS FOR PRODUCING THE ENZYME D-AMINO ACID OXIDASE  
FILE REFERENCE: U-011560-1  
CURRENT APPLICATION NUMBER: US/08/973,914F  
CURRENT FILING DATE: 1998-04-28  
EARLIER APPLICATION NUMBER: PCT/ES97/00099  
EARLIER FILING DATE: 1996-04-18  
NUMBER OF SEQ ID NOS: 9  
SEQ ID NO 9  
LENGTH: 368  
TYPE: PRT  
ORGANISM: Rhodotorula Gracilis  
US-08-973-914-9

Query Match  
Best Local Similarity 2.3%; Score 7; DB 4; Length 368;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 ETVORIL 12  
DB 250 ETVORIL 256

RESULT 8  
US-08-851-843A-52  
Sequence 52, Application US/08851843A  
Patent No. 6093809  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morlin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6093809el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 52:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-851-843A-52

Query Match  
Best Local Similarity 2.3%; Score 7; DB 3; Length 699;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TEEERYKA 146  
DB 301 TEEERYKA 307

RESULT 9  
US-08-974-549A-188  
Sequence 188, Application US/08974549A  
Patent No. 6166178  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morlin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3854  
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 188:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-974-549A-188

Query Match 2.3%; Score 7; DB 4; Length 699;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 140 TEEYKA 146  
|111111|  
Db 301 TEEYKA 307

RESULT 10  
US-08-851-843A-7  
Sequence 7, Application US/08851843A  
Patent No. 6093809  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.

APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6093809e1 Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
City: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
US-08-851-843A-7

Query Match 2.3%; Score 7; DB 3; Length 719;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 140 TEEYKA 146  
|111111|  
Db 319 TEEYKA 325

RESULT 11  
US-08-974-549A-219  
Sequence 219, Application US/08974549A  
Patent No. 6166178  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-002610US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 219:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 719 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein

US-08-974-549A-219

Query Match 2.38; Score 7; DB 4; Length 719;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 TEEYKA 146  
|||||  
Db 319 TEEYKA 325

RESULT 12  
US-08-347-792-8

Sequence 8, Application US/08347792  
Patent No. 5573925

GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 proteins with altered  
TITLE OF INVENTION: Tetramerization Domains  
NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/347,792  
FILING DATE:  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215

REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide

US-08-347-792-8

Query Match 2.08; Score 6; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 RGNPE 283  
|||||  
Db 1 RGNPE 6

RESULT 13  
US-08-431-357-8  
Sequence 8, Application US/08431357  
Patent No. 5721340

GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: p53 proteins with altered  
TITLE OF INVENTION: Tetramerization Domains  
NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/431,357  
FILING DATE:  
CLASSIFICATION: 435

US-08-431-357-8

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/347,792  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST580USA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-431-357-8

Query Match 2.0%; Score 6; DB 1; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 RGNPE 283  
|||||  
DB 1 RGNPE 6

RESULT 14  
US-08-697-221-9  
Sequence 9, Application US/08697221  
Patent No. 5847083  
GENERAL INFORMATION:  
APPLICANT: Halazonetis, Thanos D.  
TITLE OF INVENTION: Modified p53 Constructs and Uses  
TITLE OF INVENTION: Therefor  
NUMBER OF SEQUENCES: 33  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr., PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/697,221  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/004,802  
FILING DATE: 22-SEP-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Kodroff, Cathy A.  
REGISTRATION NUMBER: 33,980  
REFERENCE/DOCKET NUMBER: WST64AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9206  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-697-221-9

Query Match 2.0%; Score 6; DB 2; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 RGNPE 283  
|||||  
DB 1 RGNPE 6

RESULT 15  
US-08-920-610-6  
Sequence 6, Application US/08920610  
Patent No. 6015709  
GENERAL INFORMATION:  
APPLICANT: Natesan, Sridaran  
TITLE OF INVENTION: TRANSCRIPTIONAL ACTIVATORS, AND  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELLIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/920,610  
FILING DATE: 27-AUG-1997  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Vincent, Matthew P.  
REGISTRATION NUMBER: 36,709  
REFERENCE/DOCKET NUMBER: APV-006.02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-920-610-6

Query Match 2.0%; Score 6; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.5e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 278 RGNPE 283  
|||||  
DB 1 RGNPE 6

Search completed: August 19, 2001, 11:23:20  
Job time: 119 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 19, 2001, 11:20:36 ; Search time 12.24 Seconds  
(without alignments)  
504.665 Million cell updates/sec

Title: US-09-744-527-4  
Perfect score: 1566  
Sequence: 1 MEKQETVQRIILPEYKYL.....NPELVAVKHLKSKKEENE 300

Scoring table: BLOSUM62  
Gapop 10.0 ; Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:  
1: /cgn2.6/prodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2.6/prodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2.6/prodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2.6/prodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2.6/prodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2.6/prodata/2/iaa/backfile01.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1523	97.3	300	1	US-08-469-665-2 Sequence 2, Appl
2	1523	97.3	300	2	US-09-038-596-2 Sequence 2, Appl
3	1523	97.3	300	5	PCT-US95-00421-2 Sequence 2, Appl
4	583	37.2	335	2	US-08-761-344-2 Sequence 2, Appl
5	207	13.2	330	1	US-08-410-167A-4 Sequence 4, Appl
6	207	13.2	330	2	US-08-898-560-1 Sequence 1, Appl
7	195	12.5	325	3	US-09-009-895-2 Sequence 2, Appl
8	178.5	11.4	325	3	US-09-217-609A-2 Sequence 2, Appl
9	178.5	11.4	325	4	US-08-873-235B-2 Sequence 2, Appl
10	173	11.0	320	3	US-09-217-609A-6 Sequence 6, Appl
11	173	11.0	320	4	US-08-873-235B-6 Sequence 6, Appl
12	169	10.8	333	4	US-09-025-819-29 Sequence 29, Appl
13	156.5	10.0	297	1	US-08-534-910B-6 Sequence 6, Appl
14	155.5	9.9	297	1	US-08-534-910B-7 Sequence 7, Appl
15	155.5	9.9	297	1	US-08-534-910B-10 Sequence 10, Appl
16	154.5	9.9	297	1	US-08-534-910B-8 Sequence 8, Appl
17	154.5	9.9	297	1	US-08-534-910B-9 Sequence 9, Appl
18	154.5	9.9	297	3	US-08-886-466-2 Sequence 2, Appl
19	154.5	9.9	297	4	US-09-475-304-2 Sequence 2, Appl
20	145	9.3	293	3	US-09-276-873-2 Sequence 2, Appl
21	141.5	9.0	344	4	US-08-733-837B-6 Sequence 6, Appl
22	141	9.0	291	4	US-09-275-742-2 Sequence 2, Appl
23	124.5	8.0	285	3	US-09-187-050-12 Sequence 12, Appl
24	124.5	8.0	393	3	US-09-187-050-2 Sequence 2, Appl
25	124.5	8.0	393	3	US-09-187-050-14 Sequence 14, Appl
26	124.5	8.0	393	3	US-09-187-050-16 Sequence 16, Appl
27	124.5	8.0	393	3	US-09-187-050-18 Sequence 18, Appl

28	124.5	8.0	393	3	US-09-187-050-20 Sequence 20, Appl
29	124.5	8.0	393	3	US-09-187-050-22 Sequence 22, Appl
30	124.5	8.0	393	3	US-09-187-050-24 Sequence 24, Appl
31	124.5	8.0	393	3	US-09-187-050-26 Sequence 26, Appl
32	124.5	8.0	393	3	US-09-187-050-27 Sequence 27, Appl
33	124.5	8.0	393	3	US-09-187-050-28 Sequence 28, Appl
34	124.5	8.0	393	3	US-09-187-050-29 Sequence 29, Appl
35	124.5	8.0	393	3	US-09-187-050-30 Sequence 30, Appl
36	124.5	8.0	393	3	US-09-187-050-31 Sequence 31, Appl
37	124.5	8.0	393	3	US-09-187-050-32 Sequence 32, Appl
38	124.5	8.0	393	3	US-09-187-050-33 Sequence 33, Appl
39	124.5	8.0	393	3	US-09-187-050-34 Sequence 34, Appl
40	120	7.7	302	1	US-07-783-705A-1 Sequence 1, Appl
41	104	6.6	295	3	US-08-660-645A-1 Sequence 1, Appl
42	104	6.6	295	3	US-09-298-718-1 Sequence 1, Appl
43	104	6.6	295	4	US-09-546-969-1 Sequence 1, Appl
44	100.5	6.4	417	1	US-08-351-981-6 Sequence 6, Appl
45	100.5	6.4	417	1	US-08-351-981-7 Sequence 7, Appl

## ALIGNMENTS

RESULT 1  
US-08-469-665-2  
Sequence 2, Application US/08469665  
Patent No. 5786193  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESS: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,665  
FILING DATE: June 6, 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: 11 JAN 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-377  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS:  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-08-469-665-2  
Query Match 97.3% Score 1523; DB 1; Length 300;  
Best Local Similarity 98.3% Pred. No. 3e-145;  
Matches 295; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
QY 1 MEKQETVQRIILPEYKYLQLPKQVYRTKLSQAFNMHLKVPEDKLIITIEVTEMLNNS 60

Db 1 MEKTOEIVQRIILEPYKYLQLPKGQVRTKLSQAFNMHMLKVPEDKLOITIEVTEMLHNAS 60  
QY 61 LLIIDIEDNSKLRGFPVAHSIYIGIPSVINSANYVFLGLEKVLTLDPHDAVKLFTROLL 120  
Db 61 LLIIDIEDNSKLRGFPVAHSIYIGIPSVINSANYVFLGLEKVLTLDPHDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTPEEEKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTPEEEKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
QY 181 GLFQOIRDDYANLHKSKEYSENKSFCEDLTEGKFSFPTIHAIWSRPESTVOYONILROFTEN 240  
Db 181 GLFQOIRDDYANLHKSKEYSENKSLGEDLTEGKFSFPTIHAIWSRPESTVOYONILROFTEN 240  
QY 241 IDIKKVCVHYLEDVGSFEYTRNTLKELEAKAYKOIDARGNPDELVALVKHLSKMSKEENE 300  
Db 241 IDIKKVCVHYLEDVGSFEYTRNTLKELEAKAYKOIDARGNPDELVALVKHLSKMSKEENE 300

## RESULT 2

US-09-038-596-2  
Sequence 2, Application US/09038596  
Patent No. 5928924  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/038, 596  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/469, 665  
FILING DATE: June 6, 1995  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: 11 JAN 95  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-377  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEO ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS: LINEAR  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
US-09-038-596-2

Query Match 97.3%; Score 1523; DB 2; Length 300;  
Best Local Similarity 98.3%; Pred. No. 3e-145;

Matches 295; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEKTOEIVQRIILEPYKYLQLPKGQVRTKLSQAFNMHMLKVPEDKLOITIEVTEMLHNAS 60

Db 1 MEKTOEIVQRIILEPYKYLQLPKGQVRTKLSQAFNMHMLKVPEDKLOITIEVTEMLHNAS 60  
QY 61 LLIIDIEDNSKLRGFPVAHSIYIGIPSVINSANYVFLGLEKVLTLDPHDAVKLFTROLL 120  
Db 61 LLIIDIEDNSKLRGFPVAHSIYIGIPSVINSANYVFLGLEKVLTLDPHDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPTPEEEKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
Db 121 ELHOGGGLDIYWRDNYTCPTPEEEKAMVLOKTGGLFGLAVGLMOLFSDYKEDLKPILLNTL 180  
QY 181 GLFQOIRDDYANLHKSKEYSENKSFCEDLTEGKFSFPTIHAIWSRPESTVOYONILROFTEN 240  
Db 181 GLFQOIRDDYANLHKSKEYSENKSLGEDLTEGKFSFPTIHAIWSRPESTVOYONILROFTEN 240  
QY 241 IDIKKVCVHYLEDVGSFEYTRNTLKELEAKAYKOIDARGNPDELVALVKHLSKMSKEENE 300  
Db 241 IDIKKVCVHYLEDVGSFEYTRNTLKELEAKAYKOIDARGNPDELVALVKHLSKMSKEENE 300

## RESULT 3

PCT-US95-00421-2  
Sequence 2, Application PC/TUS9500421  
GENERAL INFORMATION:  
APPLICANT: GREENE, ET AL.  
TITLE OF INVENTION: Human Geranylgeranyl Pyrophosphate Synthetase  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,  
ADDRESSEE: CECCHI, STEWART & OLSTEIN  
STREET: 6 BECKER FARM ROAD  
CITY: ROSELAND  
STATE: NEW JERSEY  
COUNTRY: USA  
ZIP: 07068  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 INCH DISKETTE  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/00421  
FILING DATE: Concurrently  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: FERRARO, GREGORY D.  
REGISTRATION NUMBER: 36,134  
REFERENCE/DOCKET NUMBER: 325800-257  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-994-1700  
TELEFAX: 201-994-1744  
INFORMATION FOR SEO ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 300 AMINO ACIDS  
TYPE: AMINO ACID  
STRANDEDNESS: LINEAR  
TOPOLOGY: LINEAR  
MOLECULE TYPE: PROTEIN  
PCT-US95-00421-2

Query Match 97.3%; Score 1523; DB 5; Length 300;  
Best Local Similarity 98.3%; Pred. No. 3e-145;  
Matches 295; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MEKTOEIVQRIILEPYKYLQLPKGQVRTKLSQAFNMHMLKVPEDKLOITIEVTEMLHNAS 60  
Db 1 MEKTOEIVQRIILEPYKYLQLPKGQVRTKLSQAFNMHMLKVPEDKLOITIEVTEMLHNAS 60  
QY 61 LLIIDIEDNSKLRGFPVAHSIYIGIPSVINSANYVFLGLEKVLTLDPHDAVKLFTROLL 120



Db 61 LLIDIEDNSKLRGFEVASHIYIGIPVINSANVYFLGLEKVLTLDPDAVKLFTROL 120  
 QY 121 ELHOGGLDIYMRDNTCPTEEEKRANVLOKGTGLRGLANGMLQPSDYKEDKPLPNTL 180  
 Db 121 ELHOGGLDIYMRDNTCPTEEEKRANVLOKGTGLRGLANGMLQPSDYKEDKPLPNTL 180  
 QY 181 GLFPOIDDDYANLHRSKYSNKSFCEDLTGKFSFPTIHAWSRPESTOVONILRORTEN 240  
 Db 181 GLFPOIDDDYANLHRSKYSNKSFCEDLTGKFSFPTIHAWSRPESTOVONILRORTEN 240  
 QY 241 IDIKKVCVHYLEDVSGSEYTRNTLKELEAKAYOIDARGNPVELVALYKHLKMKKEENE 300  
 Db 241 IDIKKVCVHYLEDVSGSEYTRNTLKELEAKAYOIDARGNPVELVALYKHLKMKKEENE 300

RESULT 4  
 US-08-761-344-2  
 ; Sequence 2, Application US/08761344  
 ; Patent No. 5912154  
 ; GENERAL INFORMATION:

APPLICANT: Ferro-No. 5912154ick, Susan  
 APPLICANT: Jiang, Yu  
 TITLE OF INVENTION: GERANYLGERANYL DIPHOSPHATE SYNTHASE  
 TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES AND USES THEREOF  
 NUMBER OF SEQUENCES: 5  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sheridan Ross P.C.  
 STREET: 1700 Lincoln Street  
 CITY: Denver  
 STATE: CO  
 COUNTRY: U.S.A.  
 ZIP: 80203  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/761,344  
 FILING DATE:

CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 30,020  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 303/863-9700  
 TELEFAX: 303/862-0223  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 335 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-761-344-2

Query Match 37.2%; Score 583; DB 2; Length 335;  
 Best Local Similarity 40.5%; Pred. No. 1,1e-50;  
 Matches 138; Conservative 55; Mismatches 103; Indels 30; Gaps 8;

QY 9 QRLLEPKYLQLPGRQVATKLSQAFNMHLKVPEDKLOIIEVTEMLAHNASLLIDIED 68  
 Db 20 ESLSKPYNHLILKPGNFRNLILVQINRWMLPKDOLAVISOVLHLHSSLLIDIED 79  
 QY 69 NSKLRGEPVAHSIYIGIPVINSANVYFLGLEKV--LTLDHP--DAVLFTRQLLEH 123  
 Db 80 NAPIRKQTTSHLIFGVPSTINTANVIFRAMQLVSOLTKREPLYHNLITIFNEELINLH 139  
 QY 124 QGGLDIYMRD--NYTCPTTEEEKRANVLOKGTGLRGLANGMLQPSDYKEDKPLPNTL 178  
 Db 140 RGGGLDIYMRDPELLEIIPTEEMVNMVNMNTGGLFRLTLRLMEALSPSHHGHSLVPEFIN 199

QY 179 TLGIFPOIDDDYANLHRSKYSNKSFCEDLTGKFSFPTIHA-----WSRPESTOVON 232  
 Db 200 LLGIIYQIRDDYANLHRSKYSNKSFCEDLTGKFSFPTIHA-----WSRPESTOVON 259  
 QY 233 ILRORTENIDIKKVCVHYLE-DVGSFEYTRNTLKELEAKAYOIDARGNPVELVALYKHLKMKKEENE 278  
 Db 260 ILLRTSDKDKIKLQILFEDFNLSLAYTRKFNINQVNMKINDNENKYLPLDLASHSDPTAT 319  
 QY 279 GGNPELVALYKHLKMK 294  
 Db 320 NLHDELYIIDLSEL 335

RESULT 5  
 US-08-410-167A-4  
 ; Sequence 4, Application US/08410167A  
 ; Patent No. 5773273  
 ; GENERAL INFORMATION:

APPLICANT: Tokuzo NISHINO, Shinichi OHNUMA, Manabu SUZUKI,  
 APPLICANT: Chikara OHTO, Chika ASADA, Yuka HIGUCHI, Yoshie TAKEUCHI  
 TITLE OF INVENTION: Geranylgeranyl-Diphosphate Synthase and DNA  
 NUMBER OF SEQUENCES: 4  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kenyon & Kenyon  
 STREET: One Broadway  
 CITY: New York  
 STATE: NY  
 COUNTRY: US  
 ZIP: 10004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3+ Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
 SOFTWARE: Wordperfect 6.1 Windows  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/410,167A  
 FILING DATE: 24-MAR-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: JP 6-53804  
 FILING DATE: 24-MAR-1994  
 APPLICATION NUMBER: JP 6-315572  
 FILING DATE: 25-NOV-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Edward W. Greason  
 REGISTRATION NUMBER: 18,918  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212)425-7200  
 TELEFAX: (212)425-5288  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 330 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Sulfolobus acidocaldarius  
 STRAIN: ATCC 33909  
 US-08-410-167A-4

Query Match 13.2%; Score 207; DB 1; Length 330;  
 Best Local Similarity 24.7%; Pred. No. 7,1e-13;  
 Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

QY 12 LLEPKYLQLPGRQVATKLSQAFNMHLKVPEDKLOIIEVTEMLAHNASLLIDIEDNSK 71.  
 Db 30 LVEASYHLFTSGKRLPLITITSSDLFGCGRERAYAGAIEVLHFTFLVHDDIMQDN 89  
 QY 72 LRGRGEPVAHSIYIGIPVINSANVY--FLGLEKVL-LDHPDAVK--LFTQLLELHQ 124

Db 90 IRRGLPVHVKYGLPLALLAGDLHAKAFOLLTQALRGLPSETTIKARDIFTRSIITISE 149  
Oy 125 GCGGDIYWRDNYTCPTTEEEKYKAWYLQKTGGLFGLAVGL-----MOLFSDYKEDL 173  
Db 150 GOAVMEFEEDRID-KEOEYLDIMSRKTAALFSASSSIGALIGANDNDVRLMSDFG--- 205  
Oy 174 KPLNTLGLFQIRIDYANLHSEKSEKSCFEDLTGKRSFPTIHA--IWSRPESTOVQ 231  
Db 206 ----TNLGIAPQYVDDILGLTADEKELGKPVESDIRGKKTILVITLLECKEDEKRIYL 261  
Oy 232 NILRQRTENID-----IKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNP 282  
Db 262 KALGNKSASKEELMSADIIKKYSLDVAYNLAE-KYYKNALDSL-----NOVSSKSDIP 314

RESULT 6  
US-08-898-560-1  
; Sequence 1, Application us/08898560  
; Patent No. 5935832  
; GENERAL INFORMATION:  
; APPLICANT: HIROYUKI NAKANE, Chikara OHTO, Shinichi OHNUMA,  
; APPLICANT: Kazutake HIROOKA, Tokuzo NISHINO  
; TITLE OF INVENTION: Farnesyl diphosphate Synthase  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kenyon & Kenyon  
; STREET: One Broadway  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3+ Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
; SOFTWARE: WordPerfect 6.1 Windows  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/898,560  
; FILING DATE: Concurrent Herewith  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 8-213211  
; FILING DATE: 24-JUL-96  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Edward W. Greason  
; REGISTRATION NUMBER: 18,918  
; REFERENCE/DOCKET NUMBER: 77670/495  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212)425-7200  
; TELEFAX: (212)425-5288  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 330 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; ORIGINAL SOURCE:  
; ORGANISM: *Sulfolobus acidocaldarius*  
; STRAIN: ATCC 33909  
; FEATURE:  
; NAME/KEY: Asp-rich domain  
; LOCATION: 82-86  
; US-08-898-560-1

Query Match 13.2%; Score 207; DB 2; Length 330;  
Best Local Similarity 24.7%; Pred. No. 7, 1e-13;  
Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

Oy 12 LLEPYRYLLQPLGKQVTRKLSQAFNHMLVPEDKLQIIIEVEMLNANSLIDIDIEDNSK 71  
Db 30 IVEASVHLFTSGGKRRLRPLILITSSDLFGGRRRAYAAGALEVHLFTLVHDDIMDOCN 89

Oy 72 LRGEFVAHSIYGPISVINSANYV---FLGLEKVLN-LDHPAVK---LFTROLLEHQ 124  
Db 90 IRRGLPVHVKYGLPLALLAGDLHAKAFOLLTQALRGLPSETTIKARDIFTRSIITISE 149  
Oy 125 GCGGDIYWRDNYTCPTTEEEKYKAWYLQKTGGLFGLAVGL-----MOLFSDYKEDL 173  
Db 150 GOAVMEFEEDRID-KEOEYLDIMSRKTAALFSASSSIGALIGANDNDVRLMSDFG--- 205  
Oy 174 KPLNTLGLFQIRIDYANLHSEKSEKSCFEDLTGKRSFPTIHA--IWSRPESTOVQ 231  
Db 206 ----TNLGIAPQYVDDILGLTADEKELGKPVESDIRGKKTILVITLLECKEDEKRIYL 261  
Oy 232 NILRQRTENID-----IKKYCVHYLEDVGSFEYTRNTLKELEAKAYKQIDARGNP 282  
Db 262 KALGNKSASKEELMSADIIKKYSLDVAYNLAE-KYYKNALDSL-----NOVSSKSDIP 314

RESULT 7  
US-09-009-895-2  
; Sequence 2, Application US/09009895  
; Patent No. 6103488  
; GENERAL INFORMATION:  
; APPLICANT: MATSUDA, Hideyuki  
; APPLICANT: KAMAKURA, Makoto  
; APPLICANT: NAKAGAWA, Tsuyoshi  
; APPLICANT: TANAKA, Katsunori  
; TITLE OF INVENTION: METHOD OF FORMING UBIOQUINONE-10  
; FILE REFERENCE: Matsuda9  
; CURRENT APPLICATION NUMBER: US/09/009,895  
; EARLIER FILING DATE: 1998-01-21  
; EARLIER APPLICATION NUMBER: 09-244811  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 2  
; LENGTH: 325  
; TYPE: PRP  
; ORGANISM: *Rhodobacter capsulatus*  
; US-09-009-895-2

Query Match 12.5%; Score 195; DB 3; Length 325;  
Best Local Similarity 25.1%; Pred. No. 1, 1e-11;  
Matches 74; Conservative 61; Mismatches 112; Indels 48; Gaps 10;

Oy 11 LLEPYRYLLQPLGKQVTRKLSQAFNHMLVPEDKLQIIIEVEMLNANSLIDIDIEDNSK 71  
Db 34 LVMYSKHYVEAGKRMPIIMCLLAAYACGETNLKHA-----OKLAATII---EMLH 81  
Oy 58 NASLIIDIEDNSKLRGSPVAHSIYGPISVINSANYVFLGLEKVLN-LDHPAVKFLTR 117  
Db 82 TATLVHDDVDYDSGLRGRPTANATWNOTAVLVGDFLARAFLDLVDIDNNITLLKDFST 141  
Oy 118 QLELHOGGDIYWRDNYTCPTTEEEKYKAWYLQKTGGLFGLA---VGLMOLFSDYKEDL 174  
Db 142 GTCETAEQEVQLQ-AHQPTDIEDIYQIIHGKTSRLFELATEGALILACKPREYELR 200  
Oy 175 PLNTLGLFQIRIDYANLHSEKSEKSCFEDLTGKRSFPTIHAWSRPESTO--VON 232  
Db 201 RFAGHFGNAFQIIDDILYTSADPTLGRNIGDIMEGRPTPLIAAM---ONTQGEQD 256  
Oy 233 ILRQ--RTENIDIKKYCVHYLEDVGSFEY-----TRNTLKELEA-----KAYK 274  
Db 257 LIRRSIATIGTSQLEQVATAYONSALDYCHKRATEETTERALQALEIIPESTYK 311

RESULT 8  
US-09-217-609A-2  
; Sequence 2, Application US/09217609A  
; Patent No. 6071733  
; GENERAL INFORMATION:  
; APPLICANT: MURAMATSU, Masayoshi

APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tanetsoshi  
APPLICANT: SHIMIZU, Naoto  
APPLICANT: CHO, Yemwin  
TITLE OF INVENTION: Prenyl diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/217,609A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/873,235  
FILING DATE: 11-Jun-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: TOFFENETTI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-217-609A-2

Query Match 11.4%; Score 178.5; DB 3; Length 325;  
Best Local Similarity 22.7%; Pred. No. 5.1e-10;  
Matches 70; Conservative 68; Mismatches 129; Indels 41; Gaps 10;

QY 1 MEKTOETVORILLEPYKYLQLPGRQVR---TKLSQAFNHMLKVEDKQIIEVTMLH 57  
DB 25 IQSDSETINKAA---HHILSSGGRKVRPMFVLLSGFLNDQK--DDLIRAVSL-ELVH 77  
QY 58 NASLLIDIDENSKLRKRPVAHSHYIGIPSYNSANYVYFLGLEKVLVLLDHDNAKLFTR 117  
DB 78 MASLVHDYIDNSDMRKNTSVHAFDQTAIRGHFLIARALQIAIINNSKRFQISK 137  
QY 118 QLELHOGGGLDIYVRNDVYTCPTTEEYKAMVLOKTGGLFGLAVGMOLFSDVKED---L 173  
DB 138 TLEVCSEFQDMADRFNVPV-SFTAYLRIRNRKTAAILLEASCHIGALSSQDDESYTHI 196  
QY 174 KPLNTLGLFQIRDDVYANLHSEKSEKSPCEDLTEGKFSPTTHAHSRPE--STOVQ 231  
DB 197 KQFGHCIMSVQIIDDIDYTSDEATLGKPVGSDIRNCHITPTPLMAAIAANLKEQDDDKLE 256  
QY 232 NILROETENID--IKKYVHYLEDVG-----SFEY-----TRNTLKL 267  
DB 257 AVVKHLTSTSDVEYQYIVSOVKQYGIPEALLSRKYGDKAKYHLSQLODSNITKYLEEI 316  
QY 268 EAKAYKQI 275  
DB 317 HEKMLKRV 324

RESULT 9  
US-08-873-235B-2

Sequence 2, Application US/08873235B  
Patent No. 6174715  
GENERAL INFORMATION:  
APPLICANT: MURAMATSU, Masayoshi  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tanetsoshi  
APPLICANT: SHIMIZU, Naoto  
APPLICANT: CHO, Yemwin  
TITLE OF INVENTION: Prenyl diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873,235B  
FILING DATE: 11-Jun-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 154441/1996  
FILING DATE: 14-Jun-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: TOFFENETTI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 325 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-873-235B-2

Query Match 11.4%; Score 178.5; DB 4; Length 325;  
Best Local Similarity 22.7%; Pred. No. 5.1e-10;  
Matches 70; Conservative 68; Mismatches 129; Indels 41; Gaps 10;

QY 1 MEKTOETVORILLEPYKYLQLPGRQVR---TKLSQAFNHMLKVEDKQIIEVTMLH 57  
DB 25 IQSDSETINKAA---HHILSSGGRKVRPMFVLLSGFLNDQK--DDLIRAVSL-ELVH 77  
QY 58 NASLLIDIDENSKLRKRPVAHSHYIGIPSYNSANYVYFLGLEKVLVLLDHDNAKLFTR 117  
DB 78 MASLVHDYIDNSDMRKNTSVHAFDQTAIRGHFLIARALQIAIINNSKRFQISK 137  
QY 118 QLELHOGGGLDIYVRNDVYTCPTTEEYKAMVLOKTGGLFGLAVGMOLFSDVKED---L 173  
DB 138 TLEVCSEFQDMADRFNVPV-SFTAYLRIRNRKTAAILLEASCHIGALSSQDDESYTHI 196  
QY 174 KPLNTLGLFQIRDDVYANLHSEKSEKSPCEDLTEGKFSPTTHAHSRPE--STOVQ 231  
DB 197 KQFGHCIMSVQIIDDIDYTSDEATLGKPVGSDIRNCHITPTPLMAAIAANLKEQDDDKLE 256  
QY 232 NILROETENID--IKKYVHYLEDVG-----SFEY-----TRNTLKL 267  
DB 257 AVVKHLTSTSDVEYQYIVSOVKQYGIPEALLSRKYGDKAKYHLSQLODSNITKYLEEI 316  
QY 268 EAKAYKQI 275  
DB 317 HEKMLKRV 324

Sequence 6, Application US/08873235B  
Patent No. 6174715  
GENERAL INFORMATION:  
APPLICANT: MURAMATSU, Masayoshi  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OGURA, Kyoza  
APPLICANT: Koyama, Tanetoshi  
APPLICANT: SHIMIZU, Naoto  
APPLICANT: CHO, Yewlin  
TITLE OF INVENTION: Prenyl Diphosphate Synthetase Genes  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, NW - Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: US  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: WordPerfect 6.1 for Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/873.235B  
FILING DATE: 11-Jun-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 154441/1996  
FILING DATE: 14-Jun-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: TOFENETTI, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 10235/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-429-1776  
TELEFAX: 202-429-0796  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 320 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-873-235B-6

```

Query Match          11.0%; Score 173; DB 4; Length 320;
Best Local Similarity 24.1%; Pred. 0.1.8e-09;
Matches 56; Conservative 53; Mismatches 103; Indels 20; Gaps
6;

QY      5 QETVORILLEPY-----KYLIDLPKQVR--TKLSQAFNHMLKVEDKLIIEYTE 54
      ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      18 EEELERAVQSEYGLPGEAALHLLQGGRRIRRVFLLARFGQY--DLERKHVAVLE 74
      :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      55 MLIHNSLLIDDIEDNSKIRRGEPVAHSITYGLPSVINSANYTYFLGLEVYLTIDHPDAVL 114
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      75 LIHNASLIYHDVIDADARRGRPTKAAWMSNFAMTYTSDYLFARSELIRMAELGNRAQV 134
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      115 FTROLELHOGGGLDIYWRNDNYTCPE--EKKAMYLQGGFLGAVGLMOLFSPYKEDL 173
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      135 LAKTIVEVCRGEIEQ--KDKTRFQOPRLRTYLRIRRKRTALLIANSQOGLAAGAPETI 192
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

QY      174 KPLL-----NTLGLFQIORDDYANLHSEKEYSEKSCFCDLTGEGKFSPYTHAI 221
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      193 VKRLYWFCHYYGMSFQITDILDFTGTGEQJGKPRGSDLLQGNVTLPLVLYAL 244
      :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 12
US-09-025-819-29
: Sequence 29, Application us/09025819
: Patent No. 6225097
: GENERAL INFORMATION:
: APPLICANT: Obata, Shusei

```

APPLICANT: Mishino, Tokuzo  
APPLICANT: Koyama, Tanetsoshi  
APPLICANT: Sato, Yoshihiro  
TITLE OF INVENTION: DECAPRENYL DIPHOSPHATE SYNTHETASE GENE  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KENYON & KENYON  
STREET: 1500 K Street, N.W.  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/025,819  
FILING DATE: 19-FEB-1998  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 251675  
FILING DATE: 17-SEP-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Khalilian, Hourli  
REGISTRATION NUMBER: 39,546  
REFERENCE/DOCKET NUMBER: 10235/2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-220-4200  
TELEFAX: 202-220-4201  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 333 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-025-819-29

Query Match 10.8%; Score 169; DB 4; Length 333;  
Best Local Similarity: 22.9%; Pred. No. 4.8e-09;  
Matches 55; Conservative 47; Mismatches 116; Indels 22; Gaps 4;  
QY 18 YLLQPGKQVRLTSLQAFNHLKVPEDKQIIEVTEMLHNSLLIDIEDNSKLRGR 77  
DB 47 HIVEAGGKRLRPVLVLAARLCGYGNSHVLAAAVEFIHTATLDDVDVDESQRRGR 106  
QY 78 VAHSIIGISVINSANVYFELGLEKVLTLDPDAVKLFTROLLELHOGGLDIYMRDNT 137  
DB 107 TANLMDNKSVALYGVYLRARSPQLMADTESMOVMRIANASATTIAGEVLTQTLTAADVS 166  
QY 138 CPTEEYKAMVLOKTGGLAV-----GIMQFSDYKEDLKPILNTLGLFPOI 186  
DB 167 -TDEDYIIVRGKTAALSAATEAGAVVAGADPAVOQLFDG-----ALGIAFOI 218  
QY 187 RDDYANLHSEKSEKSCFEDLTEGKFSPTTHAIVSRPSTQV---QNTLQRTENIDI 243  
DB 219 VDDLDDYGGSTTIGKNVDDDFERRRLTLPVIAIARADEAERAFMERITIGORDEADL 278  
RESULT 13  
US-08-534-910B-6  
Sequence 6, Application US/08534910B  
GENERAL INFORMATION:  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OBATA, Shusei  
APPLICANT: NISHINO, Tokuzo  
APPLICANT: OHNUMA, Shinichi  
APPLICANT: NAKAZAWA, Takeshi  
APPLICANT: OGURA, Kyoze  
APPLICANT: KOYAMA, Tanetsoshi

TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable  
TITLE OF INVENTION: Of Synthesizing Geranylgeranyldiphosphate And Gene Coding T  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.  
ZIP: 20036-5405  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.25" Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: IBM/Word Perfect 6.1 Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/534,910B  
FILING DATE: 28-SEPT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-25253  
FILING DATE: 14-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Toffenetti, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 77670/398  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)429-1776  
TELEFAX: (202)429-0796  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Bacillus stearothermophilus  
US-08-534-910B-6

Query Match 10.0%; Score 156.5; DB 1; Length 297;  
Best Local Similarity: 25.5%; Pred. No. 7.3e-08;  
Matches 65; Conservative 45; Mismatches 118; Indels 27; Gaps 10;  
QY 18 YLLQPGKQVRLTSLQAFNHLKVPEDKQIIEVTEMLHNSLLIDIE--DNSKLR 74  
DB 40 YSLDAGGKRLRPILLSTVRALEKDPVAGLPVACAI-EMIHTSLIHDDLPSMDNDLRR 98  
QY 75 GPFVAHSIIGISVINSANVYFELGLEKVLTLDH---PDAVKLFTROLLE----- 121  
DB 99 GKPTNHRVGEAMAILLAGGLTYARQLTTEIDDERIPPSVRL---RLIERLAKAAGPEG 155  
QY 122 LHOGGLDIYMRDNTCPTEEYKAMVLOKTGGL--FGLAVGIMQFSDYKE--DKPIL 177  
DB 156 MWAGQAAADMEGCKTITLSELY--IHRKTGMLOYSVHAGALIGADARQIREIDERA 213  
QY 178 NTGLFPOIRDDYANLHSEKSEKSCFEDLTEGKFSPTTHAIVSRPSTQVQNTLQRT 237  
DB 214 AHGLGAFQIRDDILIDEGAEKIGKPVGSDSNKATYPALSLAGAKELAHIEAQR 273  
QY 238 -TENIDIKKCYHYL 251  
DB 274 HSRNADVGAALAYI 288  
RESULT 14  
US-08-534-910B-7  
Sequence 7, Application US/08534910B  
GENERAL INFORMATION:  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OBATA, Shusei  
APPLICANT: NISHINO, Tokuzo

APPLICANT: OHNUMA, Shinichi  
APPLICANT: NAKAZAWA, Takeshi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tanetoshi  
TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable  
TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding Ther  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.  
ZIP: 20036-5405  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.25" Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: IBM/Word Perfect 6.1 Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/534,910B  
FILING DATE: 28-SEPT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-25253  
FILING DATE: 14-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Toiffenetti, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 77670/398  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)429-1776  
TELEFAX: (202)429-0796  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Bacillus stearothermophilus  
US-08-534-910B-7

Query Match 9.9%; Score 155.5; DB 1; Length 297;  
Best Local Similarity 25.6%; Pred. No. 9.2e-08;  
Matches 66; Conservative 43; Mismatches 116; Indels 33; Gaps 11;

QY 18 YLQLPGKQVYTKL---SQAFNHWLKPEDKQIIEVTEMLHNSLLIDIE--DNSK 71  
DB 40 YSLGAGKRIRPLLLSTVQALG---KDPVAGLPVACAI-EMHTYSLIHDDLPMDND 95  
QY 72 LRGFVYASHIYIPSVNSANYVYFLGLEKVTLDH---PDAVKLFTRQLE----- 121  
DB 96 LRGRKPTNKHVFGEAMAILAGDGLTYAFQILTEIDERRIPSVRL--RLIERLAKAG 152  
QY 122 ---LHOGGGLDIYWRDNYCTPEEYKAMVLOKTGL--FGLAVGLMQLPSDYKE--DLK 174  
DB 153 PBEVMAQAADMEGKTLTLELEY--IHRHTGKMLQYSVHAGALIGADARQTRELD 210  
QY 175 PLNLTLGLFQIRDDYANLSKEYSENKSCEDLTEGKSPFTIHAIWSPSTOVONIL 234  
DB 211 EFAHNLGLAFQIRDDIIDIEGAEKIGKPVGSDOSNNKATYPALLSLAGAKKELAFHIEA 270  
QY 235 RQR-TENIDIKKYCVHYL 251  
DB 271 AQRHLRNADVGAALAYI 288

RESULT 15  
US-08-534-910B-10  
; Sequence 10, Application US/08534910B  
; Patent No. 5766911

GENERAL INFORMATION:  
APPLICANT: KOIKE, Ayumi  
APPLICANT: OBATA, Shusei  
APPLICANT: NISHINO, Tokuzo  
APPLICANT: OHNUMA, Shinichi  
APPLICANT: NAKAZAWA, Takeshi  
APPLICANT: OGURA, Kyoza  
APPLICANT: KOYAMA, Tanetoshi  
TITLE OF INVENTION: Mutated Farnesylidiphosphate Synthase Capable  
TITLE OF INVENTION: Of Synthesizing Geranylgeranylidiphosphate And Gene Coding T  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kenyon & Kenyon  
STREET: 1025 Connecticut Avenue, N.W., Suite 600  
CITY: Washington  
STATE: DC  
COUNTRY: U.S.  
ZIP: 20036-5405  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.25" Floppy Disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2  
SOFTWARE: IBM/Word Perfect 6.1 Windows  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/534,910B  
FILING DATE: 28-SEPT-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-25253  
FILING DATE: 14-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Toiffenetti, Judith L.  
REGISTRATION NUMBER: 39,048  
REFERENCE/DOCKET NUMBER: 77670/398  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)429-0796  
TELEFAX: (202)429-1776  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 297 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Bacillus stearothermophilus  
US-08-534-910B-10

Query Match 9.9%; Score 155.5; DB 1; Length 297;  
Best Local Similarity 25.6%; Pred. No. 9.2e-08;  
Matches 66; Conservative 43; Mismatches 116; Indels 33; Gaps 11;

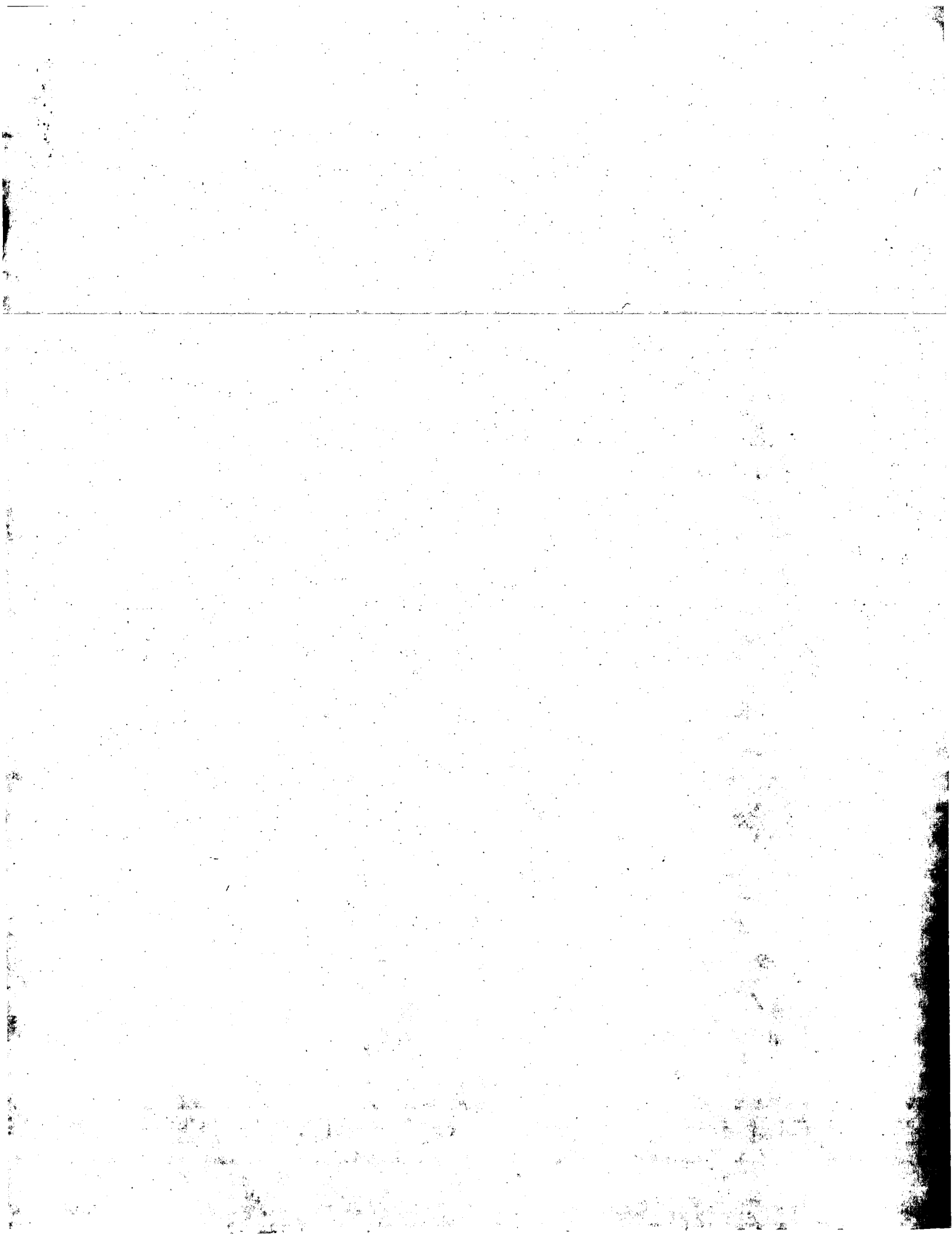
QY 18 YLQLPGKQVYTKL---SQAFNHWLKPEDKQIIEVTEMLHNSLLIDIE--DNSK 71  
DB 40 YSLGAGKRIRPLLLSTVQALG---KDPVAGLPVACAI-EMHTYSLIHDDLPMDND 95  
QY 72 LRGFVYASHIYIPSVNSANYVYFLGLEKVTLDH---PDAVKLFTRQLE----- 121  
DB 96 LRGRKPTNKHVFGEAMAILAGDGLTYAFQILTEIDERRIPSVRL--RLIERLAKAG 152  
QY 122 ---LHOGGGLDIYWRDNYCTPEEYKAMVLOKTGL--FGLAVGLMQLPSDYKE--DLK 174  
DB 153 PBEVMAQAADMEGKTLTLELEY--IHRHTGKMLQYSVHAGALIGADARQTRELD 210  
QY 175 PLNLTLGLFQIRDDYANLSKEYSENKSCEDLTEGKSPFTIHAIWSPSTOVONIL 234  
DB 211 EFAHNLGLAFQIRDDIIDIEGAEKIGKPVGSDOSNNKATYPALLSLAGAKKELAFHIEA 270  
QY 235 RQR-TENIDIKKYCVHYL 251  
DB 271 AQRHLRNADVGAALAYI 288

Mon Aug 20 11:35:43 2001

Search completed: August 19, 2001, 11:21:15  
Job time: 39 sec

us-09-744-527-4\_1.ra1

Page 9





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: August 19, 2001, 11:22:36 ; Search time 11.17 seconds

(Without alignments)  
920.022 Million cell updates/sec

Title: US-09-744-527-4

Sequence: 1 MEKQETVGRILLEPYKYL.....NPELVAVKHLKMKFRENE 300

Scoring table: OLIGO  
Gapop 60.0, Gapext 60.0

Searched: 93435 seqs, 34255486 residues

Word size: 0

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	300	100.0	300	1	GGPP_HUMAN
2	104	34.7	294	1	GGPP_BOVIN
3	70	23.3	300	1	GGPP_MOUSE
4	12	4.0	418	1	GGPP_GIBBU
5	12	4.0	428	1	GGPP_NEUCR
6	9	3.0	237	1	YKTA_BACSU
7	8	2.7	228	1	PMGY_ZYMO
8	8	2.7	672	1	ACSA_PHYBL
9	8	2.7	1790	1	USOI_YEAST
10	7	2.3	89	1	RLZ8_CHLPN
11	7	2.3	224	1	AZRL_SCHPO
12	7	2.3	231	1	PCRB_ARCFU
13	7	2.3	249	1	NEFA_BACSU
14	7	2.3	273	1	CB13_LYCES
15	7	2.3	313	1	NUIM_LOBL
16	7	2.3	314	1	NUIM_ANOGA
17	7	2.3	314	1	NUIM_ANOOU
18	7	2.3	318	1	NUIM_BRAYU
19	7	2.3	318	1	NUIM_CANFA
20	7	2.3	318	1	NUIM_FELCA
21	7	2.3	318	1	NUIM_HALGR
22	7	2.3	318	1	NUIM_MACRO
23	7	2.3	318	1	NUIM_MANTO
24	7	2.3	318	1	NUIM_MONDO
25	7	2.3	318	1	NUIM_MORTY
26	7	2.3	318	1	NUIM_NORAN
27	7	2.3	318	1	NUIM_PERGU
28	7	2.3	318	1	NUIM_PHACI
29	7	2.3	318	1	NUIM_PHOVI
30	7	2.3	318	1	NUIM_SARHA
31	7	2.3	318	1	NUIM_SMICR
32	7	2.3	318	1	NUIM_SMITA
33	7	2.3	318	1	NUIM_TAMTE

## ALIGNMENTS

RESULT	ID	GGPP_HUMAN	STANDARD	PRT	300 AA
AC	095749				
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	01-OCT-2000	(Rel. 40, Last annotation update)			
DE	GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPP SYNTHETASE)				
DE	(GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES: DIETHYLLALITYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29)]				
DE	GGPS1.				
GN	Homo sapiens (Human).				
OS	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:				
OC	Mammalia: Eutheria: Primates: Catarrhini: Homnidae: Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fetal heart;				
RX	MEDLINE=98412715; PubMed=9741684;				
RA	Ericsson J., Greene J.M., Carter K.C., Shell B.K., Duan D.R.,				
RA	Florence C., Edwards P.A.;				
RT	"Human geranylgeranyl diposphate synthase: isolation of the cDNA,				
RT	chromosomal mapping and tissue expression."				
RT	J. Lipid Res. 39:1731-1739(1998).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RX	MEDLINE=99150380; PubMed=10026212;				
RA	Kuzuguchi T., Morita Y., Sagami I., Sagami H., Ogura K.;				
RA	"Human geranylgeranyl diposphate synthase. cDNA cloning and				
RT	expression."				
RT	J. Biol. Chem. 274:5888-5894(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Misawa N., Okazaki H., Noguchi Y., Tatsuno I., Saito Y., Yasuda T.,				
RA	Hirai A.;				
RT	"Study on isolation of a geranylgeranyl pyrophosphate (GGPP) synthase				
RT	cDNA and its expression - development of a new assay system of gene				
RT	functions."				
RT	(in) Proceedings of the Japanese Conference on the Biochemistry of				
RL	Lipids, pp.41:293-296, (1999).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=pituitary;				
RX	MEDLINE=20402571; PubMed=10931946;				
RA	Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,				
RA	Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-D., Fu G., Zhang Q.-H.,				
RA	Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,				
RA	Xu S.-H.; Gu J.-J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,				
RA	Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;				
RT	"Gene expression profiling in the human hypothalamus-pituitary-adrenal				
RT	axis and full-length cDNA cloning."				
RT	Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).				
RL	[5]				
RP	SEQUENCE FROM N.A.				

078707 trichosurus  
078708 vomitus ur  
095749 aeropyrum p  
P80324 rhodosporid  
P46874 bombyx mori  
P39542 saccharomyc  
P46502 caenorhabdi  
P46507 menduca sex  
P43686 homo sapien  
P54775 mus musculu  
063570 ratius norv  
P78578 aspergillus

CC TISSUE-Liver, and Spleen;  
 RX MEDLINE=99203156; PubMed=10101267;  
 RA Kainou T., Kawamura K., Tanaka K., Matsuda H., Kawamukai M.;  
 RT Identification of the GGPS1 genes encoding geranylgeranyl diphosphate  
 RT synthases from mouse and human.  
 RL Blochim. Biophys. Acta 1437:333-340(1999).  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE, AN IMPORTANT  
 CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
 CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY.  
 CC -1- SUBUNIT: HOMOOCTAMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN TESTIS. FOUND IN OTHER  
 CC TISSUES TO A LOWER EXTENT.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AB019717; BAA75909.1; -  
 CC EMBL: AB019036; BAA77251.1; -  
 CC EMBL: AF125394; AAD3050.1; -  
 CC EMBL: AB016043; BAA76511.1; -  
 CC InterPro: IPR000092; -  
 DR Pfam: PF00348; polyprenyl\_synt.1;  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1;  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1;  
 DR Isoprene biosynthesis; transferase.  
 KW SEQUENCE 300 AA; 34871 MW; FSD1959274BEE27A CRC64;

Query Match 100.0%; Score 300; DB 1; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 1; le-295;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKTQETVORILLEPYKYLQLPKQOVRTKLSQAFNMHLKVPEDKLIITTEVTEMLHNS 60  
 DB 1 MKTQETVORILLEPYKYLQLPKQOVRTKLSQAFNMHLKVPEDKLIITTEVTEMLHNS 60  
 QY 61 LLDIEDNSKLRGFPVAHSITGIPSVINSANYVYFLGLEVLTLLDHPDAVKLTROLL 120  
 DB 61 LLDIEDNSKLRGFPVAHSITGIPSVINSANYVYFLGLEVLTLLDHPDAVKLTROLL 120  
 QY 121 ELHOGSGLDIYRDVYTCPTLEERYKAMVCKTGGLFGLAVIMOLFSYKEDKLPLNTL 180  
 DB 121 ELHOGSGLDIYRDVYTCPTLEERYKAMVCKTGGLFGLAVIMOLFSYKEDKLPLNTL 180  
 QY 181 GLEFOIRDDYAMHSKESKESKESKESKESKESKESKESKESKESKESKESKESKES 240  
 DB 181 GLEFOIRDDYAMHSKESKESKESKESKESKESKESKESKESKESKESKESKESKES 240  
 QY 241 IDIKRYCAHLEEDVGSFEETRTIKELKELAKKQIDARGNELVALVYHLSKMEKEENE 300  
 DB 241 IDIKRYCAHLEEDVGSFEETRTIKELKELAKKQIDARGNELVALVYHLSKMEKEENE 300

RESULT 2  
 ID GGP\_BOVIN STANDARD: PRT: 294 AA.  
 AC P56966;  
 DT 01-OCT-2000 (Rel. 40; Created)  
 DT 01-OCT-2000 (Rel. 40; Last sequence update)

DT 01-OCT-2000 (Rel. 40; Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPSASE)  
 DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
 DE DIMETHYLLALYLTRANSEFERASE (EC 2.5.1.1); GERANYLTRANSEFERASE  
 DE (EC 2.5.1.10); FARNESYLTRANSEFERASE (EC 2.5.1.29)] (FRAGMENT).  
 GN GGPS1.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14 AND 287-294.  
 RX MEDLINE=99150380; PubMed=10026212;  
 RA Kizuguchi T., Morita Y., Sagami H., Ogura K.;  
 RT Human geranylgeranyl diphosphate synthase. cDNA cloning and  
 RT expression.  
 RL J. Biol. Chem. 274:5888-5894(1999).  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE, AN IMPORTANT  
 CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
 CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY.  
 CC -1- SUBUNIT: HOMOOCTAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 KW Isoprene biosynthesis; Transferase.  
 FT NON\_TER 1 294  
 FT NON\_TER 1 294  
 SQ SEQUENCE 294 AA; 34138 MW; 1AC128C113FF8D53 CRC64;

Query Match 34.7%; Score 104; DB 1; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 1; le-97;  
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 KTOETVORILLEPYKYLQLPKQOVRTKLSQAFNMHLKVPEDKLIITTEVTEMLHNS 62  
 DB 1 KTOETVORILLEPYKYLQLPKQOVRTKLSQAFNMHLKVPEDKLIITTEVTEMLHNS 60  
 QY 63 IDIEDNSKLRGFPVAHSITGIPSVINSANYVYFLGLEVLTLLDHPDAVKLTROLL 106  
 DB 61 IDIEDNSKLRGFPVAHSITGIPSVINSANYVYFLGLEVLTLLDHPDAVKLTROLL 104

RESULT 3  
 ID GGP\_MOUSE STANDARD: PRT: 300 AA.  
 AC Q9WTNO;  
 DT 30-MAY-2000 (Rel. 39; Created)  
 DT 30-MAY-2000 (Rel. 39; Last sequence update)  
 DT 01-OCT-2000 (Rel. 40; Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPSASE)  
 DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
 DE DIMETHYLLALYLTRANSEFERASE (EC 2.5.1.1); GERANYLTRANSEFERASE  
 DE (EC 2.5.1.10); FARNESYLTRANSEFERASE (EC 2.5.1.29)]  
 GN GGPS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6J; TISSUE=Lymph node;  
 RX MEDLINE=99203156; PubMed=10101267;  
 RA Kainou T., Kawamura K., Tanaka K., Matsuda H., Kawamukai M.;  
 RT Identification of the GGPS1 genes encoding geranylgeranyl diphosphate  
 RT synthases from mouse and human.

RL Biochim. Biophys. Acta 1437:333-340(1999).

CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
CC IPP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE, AN IMPORTANT  
CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.

CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE - PYROPHOSPHATE + GERANYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC - PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE - PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.

CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY.

CC -1- SUBUNIT: HOMOCYMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).

CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AB016044; BAA76512.1; -  
DR MGD: MGI:1341724; Ggpsi1.  
DR InterPro: IPR000092; -  
DR Pfam: PF00348; polyprenyl\_synt.1.  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
DR Isoprene biosynthesis; Transferase.  
KW Isoprene biosynthesis; Transferase.  
SQ SEQUENCE 300 AA; 34707 MW; 13ECB67EA17EEB63 CRC64;

Query Match 23.3%; Score 70; DB 1; Length 300;  
Best Local Similarity 100.0%; Pred. No. 4.5e-63;  
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 179 TGLGFEQIRDDVYANHSKSFCEDLTEGKFSPTIHAIWSPSTOVONILRQRT 238  
Db 179 TGLGFEQIRDDVYANHSKSFCEDLTEGKFSPTIHAIWSPSTOVONILRQRT 238

Oy 239 ENIDIKRYCV 248  
Db 239 ENIDIKRYCV 248

RESULT 4  
GGPP\_GIBFU STANDARD; PRT; 418 AA.  
ID GGPP\_GIBFU STANDARD; PRT; 418 AA.  
AC 092236;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPP SYNTHETASE) (GGPPSASE)  
DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSTRANSFERASE  
DE (EC 2.5.1.10); FARNESYLTRANSTRANSFERASE (EC 2.5.1.29)].  
GN GGS OR GGPS.  
OS Gibberella fujikuroi (Fusarium moniliforme).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocerales; Nectriaceae; Gibberella.  
OX NCBI\_TaxID=5127;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M567;  
RX MEDLINE=97374453; PubMed=9230902;  
RA Mende K., Homann V., Tudzynski B.;  
RT "The geranylgeranyl diphosphate synthase gene of Gibberella fujikuroi:  
RT isolation and expression."  
RL Mol. Gen. Genet. 235:96-105(1997).  
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
CC IPP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL

CC DIPHOSPHATE - PYROPHOSPHATE + GERANYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC - PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE - PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.

CC -1- PATHWAY: BIOSYNTHESIS OF GIBBERLLINS.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X69643; CAA65644.1; -  
DR InterPro: IPR000092; -  
DR Pfam: PF00348; polyprenyl\_synt.1.  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
DR Isoprene biosynthesis; Transferase.  
FT ACT SITE 269 269 BY SIMILARITY.  
SQ SEQUENCE 418 AA; 46477 MW; BE2A0F1EA7D176C9 CRC64;

Query Match 4.0%; Score 12; DB 1; Length 418;  
Best Local Similarity 100.0%; Pred. No. 0.00027;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 205 CEDLTEGKFSFP 216  
Db 325 CEDLTEGKFSFP 336

RESULT 5  
GGPP\_NEUCR STANDARD; PRT; 428 AA.  
ID GGPP\_NEUCR STANDARD; PRT; 428 AA.  
AC P24322;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPP SYNTHETASE) (GGPPSASE)  
DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSTRANSFERASE  
DE (EC 2.5.1.10); FARNESYLTRANSTRANSFERASE (EC 2.5.1.29)].  
GN AL-3 OR ALBINO-3.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=711-OR23-1A;  
RX MEDLINE=91170267; PubMed=1826006;  
RA Caratolli A., Romano N., Ballarito P., Morelli G., Macino G.;  
RT "The Neurospora crassa carotenoid biosynthetic gene (albino 3)  
RT reveals highly conserved regions among prenyltransferases.";  
RL J. Biol. Chem. 266:5854-5859(1991).  
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
CC IPP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE - PYROPHOSPHATE + GERANYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC - PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE - PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.

CC -1- PATHWAY: CAROTENOID BIOSYNTHESIS.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- INDUCTION: BY BLUE LIGHT.

CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U20940; AAC13867.1; -  
 DR PIR: S15662; S15662.  
 DR InterPro: IPR000092; -  
 DR Pfam: PF00348; PolyPrenyl\_synth. 1.  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR Carotenoid biosynthesis; Isoprene biosynthesis; Transferase.  
 FT ACT\_SITE 268 268  
 SO SEQUENCE 428 AA; 47887 MW; 7989DEAD8E360F CRC64;

Query Match 4.0%; Score 12; DB 1; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 0 00028;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CEDTEGKFSFP 216  
 DB 322 CEDTEGKFSFP 333

RESULT 6  
 KYA\_BACSU STANDARD; PRT; 237 AA.  
 AC P21884;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE HYPOTHETICAL 27.3 KDA PROTEIN IN PDHA 5 REGION (ORF5).  
 GN KYA.  
 OS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Bacillus.  
 RX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=90368558; PubMed=1697575;  
 RA Hemila H., Palva A., Paulin L., Arvidson S., Palva I.,  
 RT "Secretory S complex of Bacillus subtilis: sequence analysis and  
 RT identity to pyruvate dehydrogenase."  
 RL J. Bacteriol. 172:5052-5063(1990).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M57435; AAA62680.1; ALT\_INIT.  
 DR Subtilisin; B610206; KYA.  
 KW Hypothetical protein.  
 SO SEQUENCE 237 AA; 27322 MW; D6924F6CB2ED724B CRC64;

Query Match 3.0%; Score 9; DB 1; Length 237;  
 Best Local Similarity 100.0%; Pred. No. 0.18;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 262 NTLKELEAK 270  
 DB 68 NTLKELEAK 76

RESULT 7  
 PMGY\_ZYMO STANDARD; PRT; 228 AA.  
 AC P30798;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE PHOSPHOGLYCERATE MUTASE (EC 5.4.2.1) (PHOSPHOGLYCEROMUTASE) (PGAM)  
 DE (BPG-DEPENDENT PGAM).  
 GN GPM OR PGM.  
 OS Zymomonas mobilis.  
 CC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;  
 CC Zymomonas.  
 RX NCBI\_TaxID=542;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-40.  
 RC STRAIN=ATCC 31821 / ZM4 / CP4;  
 RX MEDLINE=93308069; PubMed=8320209;  
 RA Yomano L.P., Scopes R.K., Ingram L.O.;  
 RT "Cloning, sequencing, and expression of the Zymomonas mobilis  
 RT phosphoglycerate mutase gene (pgm) in Escherichia coli.";  
 RL J. Bacteriol. 175:3926-3933(1993).  
 CC -1- CATALYTIC ACTIVITY: 1,3-DIPHOSPHOGLYCERATE + 3-PHOSPHOGLYCERATE =  
 CC 2,3-DIPHOSPHOGLYCERATE + 3-PHOSPHOGLYCERATE.  
 CC -1- PATHWAY: GLYCOLYSIS;  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE MUTASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L09651; AAA71937.1; -  
 DR EMBL: L09649; AAA71932.1; -  
 DR EMBL: L09650; AAA71933.1; -  
 DR PIR: C40649; C40649.  
 DR HSSP; P00950; 1BQ3.  
 DR InterPro: IPR001345; -  
 DR Pfam: PF00300; PGAM; 1.  
 DR PROSITE: PS00175; PG\_MUTASE; 1.  
 KW Isomerase; Glycolysis.  
 FT ACT\_SITE 9 9  
 FT ACT\_SITE 60 60  
 FT ACT\_SITE 179 179  
 FT ACT\_SITE 228 228  
 SO SEQUENCE 228 AA; 25938 MW; 09B7DA2610087A44 CRC64;

Query Match 2.7%; Score 8; DB 1; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 286 ALVKHLK 293  
 DB 185 ALVKHLK 192

RESULT 8  
 ACSA\_PHYBL STANDARD; PRT; 672 AA.  
 ID ACSA\_PHYBL  
 AC Q01576;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE ACETYL-COENZYME A SYNTHETASE (EC 6.2.1.1) (ACETATE-COA LIGASE) (ACYL-  
 DE ACTIVATING ENZYME).  
 GN FACA.  
 OS Phycomyces blakesleeanus.

CC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;  
 CC Phycomycetes.  
 OX NCBI\_TaxID=4837;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NRRL 1555;  
 RX MEDLINE=94335878; PubMed=7914670;  
 RA Garee V., Murillo F.J., Torres-Martinez S.;  
 RT "Isolation of the *faca* (acetyl-CoA synthetase) gene of *Phycomyces*  
 RL *blakesleeanus*.";  
 RL Mol. Gen. Genet. 244:278-286(1994).  
 CC -1- CATALYTIC ACTIVITY: ATP + ACETATE + COA = AMP + PYROPHOSPHATE +  
 CC ACETYL-COA.  
 CC -1- INDUCTION: BY ACETATE.  
 CC -1- SIMILARITY: BY OTHER ENZYMES WHICH ACT VIA AN ATP-DEPENDENT  
 CC COVALENT BINDING OF AMP TO THEIR SUBSTRATE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M94729; AAA53586.1; -  
 DR HSPSP; P08659; IIC1.  
 DR InterPro: IPR000873; -  
 DR Pfam: PF00501; AMP-binding; 1.  
 DR PROSITE: PS00455; AMP\_BINDING; 1.  
 KW Ligase.  
 SO SEQUENCE 672 AA; 74775 MW; DA6BCBA35252034E CRC64;

Query Match 2.7%; Score 8; DB 1; Length 672;  
 Best Local Similarity 100.0%; Pred. No. 4.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 99 GLEKVLTL 106  
 |||||  
 DB 590 GLEKVLTL 597

RESULT 9  
 USOL\_YEAST  
 ID USOL\_YEAST STANDARD; PRT; 1790 AA.  
 AC P25386;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE INTRACELLULAR PROTEIN TRANSPORT PROTEIN USOL.  
 GN USOL OR INT1 OR YDL058W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 CC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=X2180-1A;  
 RX MEDLINE=91185402; PubMed=2010462;  
 RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,  
 RA Yamasaki M.;  
 RT "A cytoskeleton-related gene, *usol*, is required for intracellular  
 RL protein transport in *Saccharomyces cerevisiae*.";  
 RL J. Cell Biol. 113:245-260(1991).  
 CC [2]  
 CC SEQUENCE OF 782-1790 FROM N.A.  
 RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,  
 RL Kendrick K.E.;  
 RX Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-8 FROM N.A.  
 RA Bai Y., Symington L.S.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI  
 CC COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR  
 CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE  
 CC ER AND THE GOLGI COMPLEX.  
 CC -1- DOMAIN: THE RODLIKE TAIL. SEQUENCE IS HIGHLY REPETITIVE, COMPOSED  
 CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL  
 CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.  
 CC -1- SIMILARITY: BELONGS TO THE P115(TAP)/USOL/YBL047C FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X54378; CAA38253.1; -  
 DR EMBL: L03188; AAB00143.1; -  
 DR EMBL: U53668; AAB6659.1; -  
 DR PIR: A38455; A38455.  
 DR HSPSP; P80220; IDIP.  
 DR SGD; S0002216; USOL.  
 DR Transport: Protein transport; Golgi stack; Cytoskeleton; Coiled coil;  
 KW Calcium-binding.  
 KM  
 FT DOMAIN 1 724 GLOBULAR HEAD.  
 FT DOMAIN 725 1790 COILED COIL (POTENTIAL).  
 FT DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).  
 FT DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.  
 FT DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).  
 FT CA\_BIND 475 486 POTENTIAL.  
 FT CONFLICT 847 847 G->E (IN REF. 2).  
 FT CONFLICT 924 924 E->K (IN REF. 2).  
 FT CONFLICT 1253 1253 V->I (IN REF. 2).  
 FT CONFLICT 1319 1319 I->V (IN REF. 2).  
 FT CONFLICT 1461 1461 N->S (IN REF. 2).  
 FT CONFLICT 1581 1581 G->S (IN REF. 2).  
 FT CONFLICT 1600 1600 I->V (IN REF. 2).  
 FT CONFLICT 1661 1661 R->S (IN REF. 2).  
 FT CONFLICT 1772 1772 D->DEDDDE (IN REF. 2).  
 SO SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 2.7%; Score 8; DB 1; Length 1790;  
 Best Local Similarity 100.0%; Pred. No. 11;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 173 LKPLINTL 180  
 |||||  
 DB 59 LKPLINTL 66

RESULT 10  
 RL28\_CHLPN  
 ID RL28\_CHLPN STANDARD; PRT; 89 AA.  
 AC Q928L1; Q9J067;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE 50S RIBOSOMAL PROTEIN L28.  
 GN RPBW OR RL28 OR CPN0327 OR CP0430.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 CC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=83558;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CWL029;  
 RX MEDLINE=99206606; PubMed=10192388;  
 RA Kaiman S., Mitchell W., Marathe R., Lamuel C., Fan J., Hyman R.W.,  
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
 RT "Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.";

```

RL  Nat. Genet. 21:385-389(1999).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-AR39.
RX  MEDLINE=20150255; PubMed=10684935;
RA  Read T.D., Bunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA  White O., Hickey E.K., Peterson J., Umeyan L.A., Utterback T.,
RA  Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
RA  Bowman C., Dodson R., Gwin M., Nelson W., Deboy R., Kolony J.,
RA  McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT  "Genome sequences of Chlamydia trachomatis Mohn and Chlamydia
RT  pneumoniae AR39."
RL  Nucleic Acids Res. 28:1397-1406(2000).
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN-J138.
RX  MEDLINE=20330349; PubMed=10871362;
RA  Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA  Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT  "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT  from Japan and CML029 from USA".
RL  Nucleic Acids Res. 28:2311-2314(2000).
CC  -1- SIMILARITY: BELONGS TO THE L28P FAMILY OF RIBOSOMAL PROTEINS.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AE001617; AAD18476.1; -
DR  EMBL; AF002204; AAF38272.1; -
DR  EMBL; AF002546; BAA98537.1; -
DR  TIGR; CP0430; -
DR  InterPro; IPR001383; -
DR  Pfam; PF00830; Ribosomal_L28; 1.
DR  Ribosomal protein.
KM  SEQUENCE 89 AA; 10364 MW; 7C7513DDDD16C8D8 CRC64;
SQ
Query Match 2.3%; Score 7; DB 1; Length 89;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 98 LGLEKVL 104
    |||||
DB 75 LGLEKVL 81

```

```

CC  -1- FUNCTION: CONFERS AZACYTIDINE RESISTANCE IN HIGH COPY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X98329; CAA6973.1; -
DR  SEQUENCE 224 AA; 25244 MW; 47D1AFA1C95C5F27 CRC64;
SQ
Query Match 2.3%; Score 7; DB 1; Length 224;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 105 TLDHPDA 111
    |||||
DB 43 TLDHPDA 49

```

```

RESULT 12
PCRB_ARCFU STANDARD; PRT; 231 AA.
ID PCRB_ARCFU
AC 029844;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE PCRB PROTEIN HOMOLOG.
GN PCRB OR AF0403.
OS Archaeoglobus fulgidus.
OC Archaea, Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
CC Archaeoglobus.
NX NCBI_Taxid=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwin M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kervatav A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirschner S., Delcher C.L., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Colton M.D., Spriggs T., Arlrich P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus."
RL Nature 390:364-370(1997).
CC -1- SIMILARITY: BELONGS TO THE PCRB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR  EMBL; AE001076; AAB90827.1; -
DR  TIGR; AF0403; -
DR  InterPro; IPR002911; -
DR  Pfam; PF01884; PCRB; 1.
SQ SEQUENCE 231 AA; 26144 MW; 935D2D657550ABDD CRC64;

```

```

Query Match 2.3%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 19;

```

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GNPELVA 286  
|||||  
DB 173 GNPELVA 179

RESULT 13  
NFRA\_BACSU STANDARD; PRT: 249 AA.

AC P39605: 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE NITRO/FLAVIN REDUCTASE (EC 1.-.-.-).  
GN NFRA OR NFRA1 OR IPA-43D.  
OS Bacillus subtilis.  
CC Bacteria; Firmicutes; Bacillus/Clostridium group;  
CC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=95020537; PubMed=7934828;  
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,  
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,  
RA Prescan E., Santana M., Schneider E., Schweizer J., Vertes A.,  
RA Rapoport G., Danchin A.,  
RT "Bacillus subtilis" genome project: cloning and sequencing of the 97  
RT kb region from 325 degrees to 333 degrees."  
RL Mol. Microbiol. 10:371-384(1993).  
RN [2]  
RP CHARACTERIZATION:  
RX MEDLINE=99053153; PubMed=9836433;  
RA Zemo S., Kobori T., Tanokura M., Salgo K.;  
RT "Purification and characterization of NifH, a Bacillus subtilis  
RT nitro/flavin reductase capable of interacting with the bacterial  
RT luciferase."  
RL Biosci. Biotechnol. Biochem. 62:1978-1987(1998).  
CC -1- FUNCTION: CAPABLE OF REDUCING BOTH NITROFURAZONE AND FMN  
CC EFFECTIVELY.  
CC -1- COFACTOR: FMN.  
CC -1- SIMILARITY: BELONGS TO THE FLAVIN OXIDOREDUCTASE FRP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: X73124; CASI599.1;  
DR EMBL: 299123; CAB15837.1;  
DR PIR: S39698; S39698.  
DR HSSP: 056691; 1BKJ.  
DR Subtilist; BG10589; nfrA.  
DR InterPro: IPR000415;  
DR Pfam: PF00881; Nitroreductase; 1.  
KW Oxidoreductase; Flavoprotein; FMN.  
SQ SEQUENCE 249 AA; 28320 MW; A36FDAEC3692CF CRC64;

Query Match 2.3%; Score 7; DB 1; Length 249;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 155 LFGLAVG 161  
|||||  
DB 158 LFGLAVG 164

RESULT 14

CB13\_LYCES  
ID CB13\_LYCES STANDARD; PRT: 273 AA.

AC P27522: 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CHLOROPHYLL A-B BINDING PROTEIN 8 PRECURSOR (LMCI TYPE III CAB-8).  
GN CAB8.  
OS Lycopersicon esculentum (Tomato).  
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
CC Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I;  
CC Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4081;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pichersky E., Brock T.G., Nguyen D., Hoffman N.E., Piechulla B.,  
RA Tanksley S.D., Green B.R.;  
RT "A new member of the CAB gene family: structure, expression and  
RT chromosomal location of cab-8, the tomato gene encoding the type III  
RT chlorophyll a/b-binding polypeptide of photosystem I.";  
RL Plant Mol. Biol. 12:257-270(1989).  
CC -1- FUNCTION: THE LIGHT-HARVESTING COMPLEX (LHC) FUNCTIONS AS A LIGHT  
CC RECEPTOR, IT CAPTURES & DELIVERS EXCITATION ENERGY TO PHOTOSYSTEMS  
CC WITH WHICH IT IS CLOSELY ASSOCIATED. THE N-TERMINUS OF THE PROTEIN  
CC EXTENDS INTO THE STROMA WHERE IT IS INVOLVED WITH ADHESION OF  
CC GRANA MEMBRANES AND PHOTOREGULATED BY REVERSIBLE PHOSPHORYLATION  
CC OF ITS THREONINE RESIDUES; BOTH ARE BELIEVED TO MEDIATE THE  
CC DISTRIBUTION OF EXCITATION ENERGY BETWEEN PHOTOSYSTEMS I AND II.  
CC -1- SUBUNIT: THE LHC COMPLEX CONSISTS OF CHLOROPHYLLS (A & B) AND  
CC CHLOROPHYLL A-B BINDING PROTEINS.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST THYLAKOID MEMBRANE.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: X15258; CAA33330.1;  
DR PIR: S04125; S04125.  
DR InterPro: IPR001344;  
DR Pfam: PF00504; chloro.a-b-bind; 1.  
KW Chlorophyll; Photosynthesis; Photosystem I; Photosystem II;  
KW Thylakoid membrane; Chloroplast; Transit peptide; Multigene family;  
KW Transmembrane; Phosphorylation.  
FT TRANSIT 1 32 CHLOROPLAST (POTENTIAL).  
FT CHAIN 33 273 CHLOROPHYLL A-B BINDING PROTEIN 8.  
FT TRANSMEM 231 247 POTENTIAL.  
SQ SEQUENCE 273 AA; 29362 MW; 5F62013320620DBC CRC64;

Query Match 2.3%; Score 7; DB 1; Length 273;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 96 YFLGLEK 102  
|||||  
DB 184 YFLGLEK 190

RESULT 15  
NUIW\_LOLBL  
ID NUIW\_LOLBL STANDARD; PRT: 313 AA.

AC O47479;  
DT 01-OCT-2000 (Rel. 40, Created)  
DT 01-OCT-2000 (Rel. 40, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN I (EC 1.6.5.3).  
GN NDI.  
OS Loligo bleekeri (Bleeker's squid).

OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea;  
 OC Myopsida; Loliginidae; Loligo.  
 OX NCBI\_TaxID=6617;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Tomita K., Ueda T., Watanabe K.;  
 RT "Completing of squid (Loligo brekeri) mitochondrial genome  
 sequencing".  
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SIMILARITY: BELONGS TO THE COMPLEX I SUBUNIT 1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC  
 CC EMBL; AB009838; BA24062.1;  
 DR EMBL; AB029616; BAB03649.1;  
 DR InterPro; IPR001694;  
 DR Pfam; PF00146; NADHdh; 1.  
 DR PROSITE; PS00667; COMPLEX1\_ND1\_1; FALSE\_NEG.  
 DR PROSITE; PS00668; COMPLEX1\_ND1\_2; 1.  
 KW Oxidoreductase; NAD; ubiquinone; Mitochondrion; Transmembrane.  
 KW  
 SQ SEQUENCE 313 AA; 34914 MW; FB3E151B8C710CF CRC64;

Query Match 2.3%; Score 7; DB 1; Length 313;  
 Best local Similarity 100.0%; Pred. No. 24;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 110 DAVKFT 116  
 DB 54 DAVKFT 60

Search completed: August 19, 2001, 11:24:30  
 Job time: 114 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2001, 11:21:41 : Search time 15.98 Seconds  
(without alignments)  
1430.060 Million cell updates/sec

Title: US-09-744-527-4

Sequence: 1 MEKTOETVQRILPEPKYLL.....NPELVALVKHLSKMKKEENE 300

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 219241 seqs, 76174552 residues

Word size : 0

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

PIR-68: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	12	4.0	428	2	farnesyltransferase
2	10	3.3	335	2	farnesyltransferase
3	9	3.0	205	2	conserved hypothet
4	8	2.7	228	2	phosphoglycerate m
5	8	2.7	277	2	hypothetical prote
6	8	2.7	347	2	octaprenyl-diphosp
7	8	2.7	606	2	hypothetical prote
8	8	2.7	672	2	acetate--CoA ligas
9	8	2.7	1790	2	transporth protein
10	7	2.3	62	2	hypothetical prote
11	7	2.3	89	2	ribosomal protein
12	7	2.3	89	2	128 ribosomal prot
13	7	2.3	105	2	hypothetical prote
14	7	2.3	174	2	hypothetical prote
15	7	2.3	183	2	anaerobic ribonuc
16	7	2.3	231	2	conserved hypothet
17	7	2.3	247	2	oxidoreductase, sh
18	7	2.3	249	2	nitro/flavin reduc
19	7	2.3	273	2	chlorophyll a/b-bi
20	7	2.3	273	2	PSI type III chl
21	7	2.3	275	2	probable chlorophy
22	7	2.3	280	2	conserved hypothet
23	7	2.3	286	2	chlorophyll a/b-bi
24	7	2.3	287	2	probable glutamate
25	7	2.3	288	2	azrl protein [impo
26	7	2.3	293	2	methylviologen-red
27	7	2.3	299	2	hypothetical prote
28	7	2.3	314	2	NADH dehydrogenase
29	7	2.3	314	2	NADH dehydrogenase

30	7	2.3	318	2	NADH dehydrogenase
31	7	2.3	318	2	NADH dehydrogenase
32	7	2.3	318	2	NADH dehydrogenase
33	7	2.3	318	2	NADH dehydrogenase
34	7	2.3	318	2	NADH dehydrogenase
35	7	2.3	318	2	NADH dehydrogenase
36	7	2.3	318	2	NADH dehydrogenase
37	7	2.3	318	2	NADH dehydrogenase
38	7	2.3	318	2	NADH dehydrogenase
39	7	2.3	318	2	NADH dehydrogenase
40	7	2.3	318	2	NADH dehydrogenase
41	7	2.3	318	2	NADH dehydrogenase
42	7	2.3	318	2	NADH dehydrogenase
43	7	2.3	318	2	NADH dehydrogenase
44	7	2.3	318	2	NADH dehydrogenase
45	7	2.3	318	2	NADH dehydrogenase

#### ALIGNMENTS

RESULT 1 102 (b)

S15662 farnesyltransferase (EC 2.5.1.29) [validated] - Neurospora crassa  
N: Alternate names: geranylgeranyl pyrophosphate synthetase  
C: Species: Neurospora crassa  
C: Date: 21-Nov-1993 #sequence, revision 10-Nov-1995 #text, change 02-Sep-2000  
C: Accession: S15662; T46596; T46592  
R: Carattoli, A.; Romano, N.; Ballarín, P.; Morelli, G.; Macino, G.  
J. Biol. Chem. 266, 5854-5859, 1991  
A: Title: The Neurospora crassa carotenoid biosynthetic gene (albino 3) reveals highly  
A: Reference number: S15662; MUID: 91170267  
A: Accession: S15662  
A: Status: preliminary  
A: Molecule type: DNA  
A: Residues: 1-428 <CAR>  
A: Cross-references: GB:U20940; EMBL:X53979; NID:9903318; PIDN:AMC13867.1; PID:9903319.  
R: Vittorioso, P.; Carattoli, A.; Londel, P.; Macino, G.  
J. Biol. Chem. 269, 26650-26654, 1994  
A: Title: Internal translational initiation in the mRNA from the Neurospora crassa alb  
A: Reference number: A55065; MUID: 95014519  
A: Accession: T46598  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 'MEHT', 1-129 <VIT1>  
A: Cross-references: EMBL:S74011; PIDN: CAB33185.1  
A: Accession: T46592  
A: Status: preliminary; translated from GB/EMBL/DBJ  
A: Molecule type: DNA  
A: Residues: 'MEHT', 1-77, 'PACGSLPRRHQWRSPSLPTQTDLPK', 'TSSARPGAPKRRRRF' <VIT2>  
A: Cross-references: EMBL:S74683; PIDN: CAB33252.1  
A: Experimental source: mutant rosy  
A: Note: translation starts at an alternate initiator and produces a partially active  
C: Genetics:  
A: Gene: al-3  
C: Keywords: transferase

Query Match 4.0%; Score 12; DB 2; Length 428;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 205 CEDLTGKFSFP 216 V  
DB 322 CEDLTGKFSFP 333

RESULT 2

S60921 farnesyltransferase (EC 2.5.1.29) - yeast (Saccharomyces cerevisiae).  
N: Alternate names: geranylgeranyl-diphosphate synthase; protein lpe1c; protein YPL069  
C: Species: Saccharomyces cerevisiae  
C: Date: 15-Feb-1996 #sequence, revision 01-Mar-1996 #text, change 24-Nov-1999

C:Accession: S60921; S61936  
 R:Winnett, E.; Ahmed, A.; Bussey, H.; Portin, N.; Friesen, J.D.; Hall, J.; Storms, R.K.;  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: S60921  
 A:Accession: S60921  
 A:Molecule type: DNA  
 A:Residues: 1-335 <MIN>  
 A:Cross-references: EMBL:U09205; NID:g1079672; PIDN:AA68296.1; PID:g1079673; MIPS:YPL06  
 R:Jiang, Y.; Proteau, P.; Poulter, D.; Ferro-Novick, S.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: Bts1 encodes a geranylgeranyl diphosphate synthase in *Saccharomyces cerev*  
 A:Reference number: S61936  
 A:Accession: S61936  
 A:Molecule type: DNA  
 A:Residues: 1-335 <TIA>  
 A:Cross-references: EMBL:U031632; NID:g1098640; PIDN:AAA83262.1; PID:g1098641  
 C:Genetics:  
 A:Gene: SGD:BTS1  
 A:Cross-references: SGD:S0005990; MIPS:YPL069c  
 A:Map position: 16L  
 C:Superfamily: pren1 transferase A  
 C:Keywords: transferase

Query Match 3.3%; Score 10; DB 2; Length 335;  
 Best Local Similarity 100.0%; Pred. No. 0.036;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 SLIDIEDN 69  
 |||||  
 DB 71 SLIDIEDN 80

RESULT 3  
 D69870  
 conserved hypothetical protein ykya - *Bacillus subtilis*  
 N:Alternate names: hypothetical protein (acea 5' region)  
 C:Species: *Bacillus subtilis*  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999  
 C:Accession: D69870; A36718  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berlet  
 C.; Brown, S.; Brouillet, S.; Brusch, C.V.; Caldwel, B.; Capuano, V.; Carter, N.M.; Chn  
 A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 iech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
 Y, M.; Ogawa, K.; Ogawara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell  
 Rieger, M.; Rivoita, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schaefer, R.; Scottone, F.; Sekiguchi, J.; Sekowska, A.; Senot  
 akouchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
 T.; Whalers, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A:Reference number: A6580; MUID:98044033  
 A:Accession: D69870  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-205 <KUN>  
 A:Cross-references: GB:AL009126; NID:G2633699; PIDN:CAB13330.1; PID:ell185047;  
 A:Experimental source: strain 168  
 R:Hemilae, H.; Palva, A.; Paulin, L.; Arvidson, S.; Palva, I.  
 J. Bacteriol. 172, 5052-5063, 1990  
 A>Title: Secretory S complex of *Bacillus subtilis*: sequence analysis and identity to pY  
 A:Reference number: A36718; MUID:90368558  
 A:Accession: A36718  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 186-205 <HEM>  
 A:Cross-references: GB:M31542  
 C:Genetics:  
 A:Gene: ykya

Query Match 3.0%; Score 9; DB 2; Length 205;  
 Best Local Similarity 100.0%; Pred. No. 0.26;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 262 NTLKELEAK 270  
 |||||  
 DB 36 NTLKELEAK 44

RESULT 4  
 C40649  
 phosphoglycerate mutase (EC 5.4.2.1) - *Zymomonas mobilis*  
 C:Species: *Zymomonas mobilis*  
 C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 16-Jul-1999  
 C:Accession: C40649  
 R:Yomano, L.P.; Scopes, R.K.; Ingram, L.O.  
 J. Bacteriol. 175, 3926-3933, 1993  
 A>Title: Cloning, sequencing, and expression of the *Zymomonas mobilis* phosphoglycerat  
 A:Reference number: A40649; MUID:93308069  
 A:Accession: C40649  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-228 <YOM>  
 A:Cross-references: GB:I09651; NID:G155609; PIDN:AAA71937.1; PID:G155611  
 C:Superfamily: phosphoglycerate mutase; phosphoglycerate mutase homology  
 C:Keywords: intramolecular transferase; isomerase; phosphohistidine; phosphoprotein  
 F:4-214/Domain: phosphoglycerate mutase homology <PKM>  
 F:9/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 2.7%; Score 8; DB 2; Length 228;  
 Best Local Similarity 100.0%; Pred. No. 3.1;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 ALVKHLISK 293  
 |||||  
 DB 185 ALVKHLISK 192

RESULT 5  
 T47153  
 hypothetical protein DKFZp564A0772.1 - human  
 C:Species: *Homo sapiens* (man)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47153  
 R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Well, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24379  
 A:Accession: T47153  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-277 <AAA>  
 A:Cross-references: EMBL:AL161994  
 A:Experimental source: fetal brain; clone DKFZp564A0772  
 C:Genetics:  
 A>Note: DKFZp564A0772.1

Query Match 2.7%; Score 8; DB 2; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 3.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 142 BEYKAMVL 149  
 |||||  
 DB 70 BEYKAMVL 77

RESULT 6  
 A70139  
 octaprenyl-diphosphate synthase (Ispr) homolog - Lyme disease spirochete  
 C:Species: *Borrelia burgdorferi* (Lyme disease spirochete)

C>Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999  
 C/Accession: A70139  
 R:Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White  
 son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salberg, S.; Hanson, M.; Vugt,  
 Bowman, C.; Garland, S.; Fujili, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
 Nature 390, 580-586, 1997  
 A:Authors: Smith, H.O.; Venter, J.C.  
 A>Title: Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*.  
 A:Reference number: A70100; MUID:98065943  
 A/Accession: A70139  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A/Molecule type: DNA  
 A:Residues: 1-347 <KDE>  
 A/Cross-references: GB:AE001138; GB:AE000783; NID:92688210; PIDN:AAC66696.1; PID:9268821  
 A/Experimental source: strain B31  
 C:Superfamily: prenyl transferase A

Query Match 2.7%; Score 8; DB 2; Length 347;  
 Best Local Similarity 100.0%; Pred. No. 4.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 63 IIDIEDNS 70  
 Db 100 IIDIEDNS 107

RESULT 7  
 T29190  
 hypothetical protein C55C3.1 - *Caenorhabditis elegans*  
 C/Species: *Caenorhabditis elegans*  
 C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 R:Moessne, J.; Stellyes, L.  
 submitted to the EMBL Data Library, April 1996  
 A/Description: The sequence of C. elegans cosmid C55C3.  
 A/Reference number: 220585  
 A/Accession: T29190  
 A>Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A:Residues: 1-606 <MOE>  
 A/Cross-references: EMBL:U53335; PIDN:AAA96173.1; GSPDB:GN00022; CESP:C55C3.1  
 A/Experimental source: strain Bristol N2; clone C55C3  
 C/Genetics:  
 A:Gene: CESP:C55C3.1  
 A/Map position: 4  
 A/Introns: 61/3; 99/2; 126/2; 161/3; 283/2; 359/2; 424/1; 463/3; 522/1; 570/1

Query Match 2.7%; Score 8; DB 2; Length 606;  
 Best Local Similarity 100.0%; Pred. No. 7.5;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 153 GGLFGLAV 160  
 Db 191 GGLFGLAV 198

RESULT 8  
 S46276  
 acetate--CoA ligase (EC 6.2.1.1) - *Phycomyces blakesleeanus*  
 N:Alternate names: acetyl-coenzyme A synthetase  
 C/Species: *Phycomyces blakesleeanus*  
 C/Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 05-May-2000  
 C/Accession: S46276  
 R:Garre, V.; Murillo, F.J.; Torres-Martinez, S.  
 Mol. Gen. Genet. 244, 278-286, 1994  
 A>Title: Isolation of the *faca* (acetyl-CoA synthetase) gene of *Phycomyces blakesleeanus*.  
 A/Reference number: S46276; MUID:94335878  
 A/Accession: S46276  
 A>Status: preliminary  
 A/Molecule type: DNA  
 A:Residues: 1-672 <GAR>

A/Cross-references: EMBL:M94729; NID:9169007; PIDN:AAA53586.1; PID:9169008  
 C/Genetics:  
 A:Introns: 56/3; 534/1; 604/1; 652/2  
 C:Superfamily: acetate--CoA ligase; acetate--CoA ligase homology  
 C/Keywords: acid-thiol ligase; coenzyme A  
 F:142-631/Domain: acetate--CoA ligase homology <ACLU>

Query Match 2.7%; Score 8; DB 2; Length 672;  
 Best Local Similarity 100.0%; Pred. No. 8.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 99 GLEKVLTL 106  
 Db 590 GLEKVLTL 597

RESULT 9  
 S67593  
 transport protein USO1 - yeast (*Saccharomyces cerevisiae*)  
 N:Alternate names: protein D2552; protein YDL058w  
 C/Species: *Saccharomyces cerevisiae*  
 C/Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 21-Jul-2000  
 C/Accession: S67593; A38455; S30782  
 R:Blöcker, H.; Brandt, P.  
 submitted to the Protein Sequence Database, July 1996  
 A/Reference number: S67587  
 A/Accession: S67593  
 A/Molecule type: DNA  
 A:Residues: 1-1790 <BLO>  
 A/Cross-references: EMBL:Z74106; NID:91431058; PID:e253003; PID:91431059; MIPS:YDL058  
 A/Experimental source: strain S288C  
 R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.  
 J. Cell Biol. 113, 245-260, 1991  
 A>Title: A cytoskeleton-related gene, USO1, is required for intracellular protein tra  
 A/Reference number: A38455; MUID:91185402  
 A/Accession: A38455  
 A/Molecule type: DNA  
 A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAC>  
 A/Cross-references: GB:X54378; NID:94777; PIDN:CAA38253.1; PID:94778  
 A/Note: the authors translated the codon ACT for residue 768 as Ile  
 R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.  
 submitted to the EMBL Data Library, February 1993  
 A/Description: An integrin analogue in *Saccharomyces cerevisiae*.  
 A/Reference number: S30782  
 A/Accession: S30782  
 A/Molecule type: DNA  
 A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580  
 A/Cross-references: EMBL:L03188  
 C/Genetics:  
 A:Gene: SGD:USO1; INT1  
 A/Cross-references: SGD:S0002216; MIPS:YDL058w  
 A/Map position: 4L  
 C/Keywords: coiled coil; transmembrane protein  
 F:326-342/Domain: transmembrane #status predicted <TM1>  
 F:394-410/Domain: transmembrane #status predicted <TM2>  
 F:617-633/Domain: transmembrane #status predicted <TM3>

Query Match 2.7%; Score 8; DB 2; Length 1790;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 173 LKPLINTL 180  
 Db 59 LKPLINTL 66

RESULT 10  
 T36447  
 hypothetical protein SCF43A.24c - *Streptomyces coelicolor*  
 C/Species: *Streptomyces coelicolor*  
 C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T36447  
R:Seeger, K.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, July 1999  
A:Reference number: Z21598  
A:Accession: T36447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-62 <STO>  
A:Cross-references: EMBL:AL096837; PIDN:CAB48911.1; GSPDB:GN00070; SCOEDB:SCF43A.24C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCF43A.24C

Query Match 2.3%; Score 7; DB 2; Length 62;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 237 RTENIDI 243  
|||||  
DB 13 RTENIDI 19

## RESULT 11

ribosomal protein L28 CP0430 [imported] - Chlamydia pneumoniae (strains CWL029 and A  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Nov-2000

C:Accession: H72091; E81577  
R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Ollinger, L.; Grimwood, J.  
Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: H72091

A:Molecule type: DNA  
A:Residues: 1-89 <ARN>  
A:Cross-references: GB:AE001617; GB:AE001363; NID:94376599; PIDN:AD18476.1; PID:9437660

A:Experimental source: strain CWL029  
R:Head, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MOpn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255  
A:Accession: E81577

A:Molecule type: DNA  
A:Residues: 1-89 <REA>  
A:Cross-references: GB:AE002204; GB:AE002161; NID:97189350; PIDN:AAF38272.1; PID:9718935

A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: r128; CP0430  
C:Superfamily: Escherichia coli ribosomal protein L28

Query Match 2.3%; Score 7; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 LGLEKVL 104  
|||||  
DB 75 LGLEKVL 81

## RESULT 12

G86531  
L28 ribosomal protein [imported] - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001

C:Accession: G86531  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; IS

Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349  
A:Accession: G86531

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-89 <STO>  
A:Cross-references: GB:BA000008; NID:98978701; PIDN:BA98537.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: r128

Query Match 2.3%; Score 7; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 LGLEKVL 104  
|||||  
DB 75 LGLEKVL 81

## RESULT 13

T24359  
hypothetical protein T02E1.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T24359  
R:Jennard, N.

Submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19881  
A:Accession: T24359

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-105 <WILD>

A:Cross-references: EMBL:Z81581; PIDN:CAB04660.1; GSPDB:GN00019; CESP:T02E1.4  
A:Experimental source: clone T02E1  
C:Genetics:  
A:Gene: CESP:T02E1.4  
A:Map position: 1

Query Match 2.3%; Score 7; DB 2; Length 105;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 263 TLKLEA 269  
|||||  
DB 81 TLKLEA 87

## RESULT 14

G83712  
hypothetical protein BH0503 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000

C:Accession: G83712  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; H

Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a

A:Reference number: A83650; MUID:20263314  
A:Accession: G83712

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-174 <STO>

A:Cross-references: GB:AP001508; GB:BA000004; NID:910172890; PIDN:BA04222.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH0503  
C:Superfamily: Deinococcus radiodurans hypothetical protein DR0763

Query Match 2.3%; Score 7; DB 2; Length 174;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 LGLEKVL 104

Db 127 LGLEKVL 133

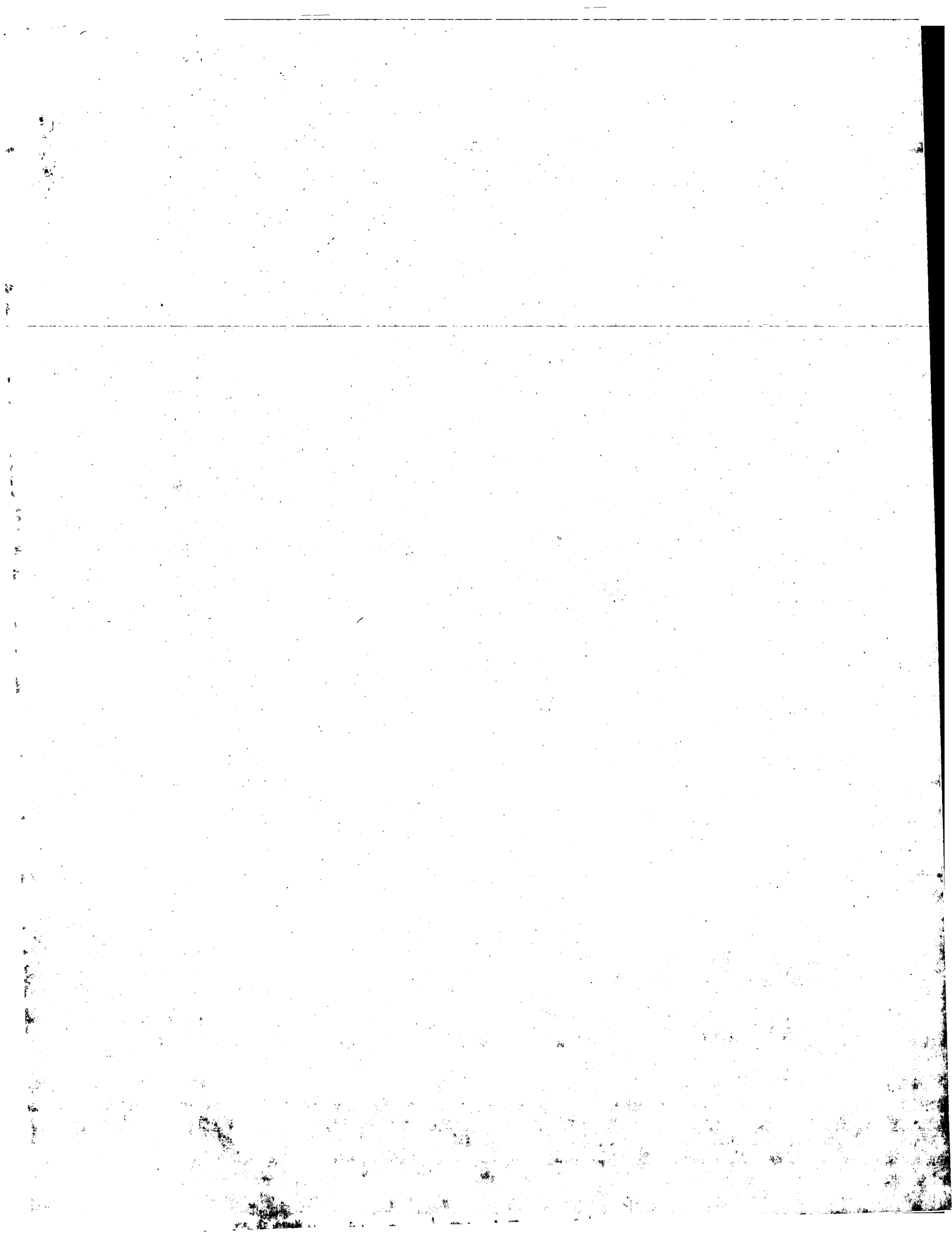
RESULT 15

H72383 anaerobic ribonucleoside-triphosphate reductase-related protein - Thermotoga maritima (S  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: H72383  
R:Neelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq  
A:Reference number: A72200; MUID:99287316  
A:Accession: H72383  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-183 <ARN>  
A:Cross-references: GB:AE001718; GB:AE000512; NID:g4980881; PID:AD35469.1; PID:g498088  
A:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM0384

Query Match 2.38; Score 7; DB 2; Length 183;  
Best Local Similarity 100.08; Pred. No. 28;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 291 LSKMFKE 297  
|||  
Db 92 LSKMFKE 98

Search completed: August 19, 2001, 11:23:43  
Job time: 122 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2001, 11:20:36 ; Search time 11.26 Seconds  
(without alignments)  
912.668 Million cell updates/sec

Title: US-09-744-527-4  
Perfect score: 1566  
Sequence: 1 MEKTOETVQRIILEPKYL.....NPELVAVKHLKMKKEENE 300

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt-39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1566	100.0	300	1	GGPP_HUMAN
2	1516	96.8	294	1	GGPP_BOVIN
3	1476	94.3	300	1	GGPP_MOUSE
4	797.5	50.9	428	1	GGPP_NEUCR
5	754.5	48.2	418	1	GGPP_GIBFU
6	292	18.6	327	1	GGPP_GIBFU
7	233.5	14.9	324	1	IDSA_METTA
8	231.5	14.8	325	1	IDSA_METTM
9	217	13.9	323	1	PREA_CYPAP
10	209.5	13.4	332	1	GGPP_SULSO
11	208.5	13.3	323	1	ISPB_ECOLI
12	207	13.2	330	1	GGPP_ECOLI
13	195	12.5	348	1	HEP2_BACSU
14	182.5	11.7	329	1	ISPB_HAEIN
15	173	11.0	320	1	HEP2_BACST
16	163.5	10.4	295	1	ISPB_HAEIN
17	163.5	10.4	323	1	PREA_PORPU
18	159	10.2	272	1	ISPB_BACSU
19	158	10.1	300	1	CRTE_CYPAP
20	154.5	9.9	297	1	ISPB_BACST
21	150.5	9.6	369	1	GGPP_CAPAN
22	148	9.5	299	1	ISPB_ECOLI
23	142	9.1	371	1	GGPP_ARATH
24	141.5	9.0	262	1	GGPP_AQUAE
25	140.5	9.0	347	1	GGPP_SCHPO
26	137	8.7	353	1	GGPP_HUMAN
27	136.5	8.7	347	1	GGPP_GIBFU
28	133.5	8.5	291	1	ISPB_MICLU
29	132.5	8.5	353	1	GGPP_RAT
30	130.5	8.3	347	1	GGPP_NEUCR
31	129	8.2	359	1	GGPP_MYCTO
32	128.5	8.2	349	1	GGPP_KLULA
33	127	8.1	288	1	CRTE_RHOSH

## ALIGNMENTS

RESULT ID	1	GGPP_HUMAN	STANDARD:	PRT:	300 AA.
AC	095749;				
DT	30-MAY-2000 (Rel. 39, Created)				
DT	01-OCT-2000 (Rel. 40, Last sequence update)				
DE	GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPP SYNTHETASE) (GGPPSASE)				
DE	(GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES: DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29)]				
GN	GGPSI				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Fetal heart;				
RX	MEDLINE=98412715; PubMed=9741684;				
RA	Ericsson J., Greene J.M., Carter K.C., Shell B.K., Duan D.R.,				
RA	Florence C., Edwards P.A.;				
RT	"Human geranylgeranyl diphosphate synthase: Isolation of the cDNA,				
RT	chromosomal mapping and tissue expression.";				
RT	J. Lipid Res. 39:1731-1739(1998).				
RL	[2]				
RL	SEQUENCE FROM N.A.				
RC	TISSUE=Testis;				
RX	MEDLINE=99150380; PubMed=10026212;				
RA	Kuzunuchi T., Morita Y., Sagami I., Sagami H., Ogura K.;				
RA	"Human geranylgeranyl diphosphate synthase. cDNA cloning and				
RT	expression.";				
RT	J. Biol. Chem. 274:5888-5894(1999).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RA	Misawa N., Okazaki H., Noguchi Y., Tatsuno I., Saito Y., Yasuda T.,				
RA	Hirai A.;				
RT	"Study on isolation of a geranylgeranyl pyrophosphate (GGPP) synthase				
RT	cDNA and its expression - development of a new assay system of gene				
RT	functions.";				
RT	(In) Proceedings of the Japanese Conference on the Biochemistry of				
RT	Lipids, pp.41:293-296, (1999).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Plutary;				
RX	MEDLINE=20402571; PubMed=10931946;				
RA	Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,				
RA	Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,				
RA	Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,				
RA	Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;				
RT	"Gene expression profiling in the human hypothalamus-pituitary-adrenal				
RT	axis and full-length cDNA cloning.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).				
RN	[5]				
RP	SEQUENCE FROM N.A.				

34	126.5	8.1	367	1	FPPS_CHICK	P08836 gallus gall
35	122.5	7.8	282	1	ISPB_BUCAL	P57537 buchnera ap
36	121	7.7	343	1	FPPS_ARATH	O09152 arabidopsis
37	120	7.7	289	1	CRTE_RHOCA	P17060 rhodobacter
38	120	7.7	302	1	CRTE_PANAN	P21684 pantoea ana
39	119.5	7.6	352	1	FPPS_YEAST	P08524 saccharomyc
40	115.5	7.4	473	1	CCO1_YEAST	P18900 saccharomyc
41	114	7.3	299	1	COMO_BACSU	P33690 bacillus su
42	114	7.3	342	1	FPP1_LUPAL	P49351 lupinus alb
43	114	7.3	429	1	KICK_HUMAN	P05783 homo sapien
44	110	7.0	342	1	FPP2_LUPAL	P49352 lupinus alb
45	107	6.8	350	1	FPPS_MAIZE	P49353 zea mays (m

CC TISSUE: liver, and Spleen:  
 RX MEDLINE=99203156; PubMed=10101267; Matsuda H., Kawamukai M.;  
 RA Kainou T., Kawamura K., Tanaka K.,  
 RT "Identification of the GGPST genes encoding geranylgeranyl diphosphate  
 RT synthases from mouse and human."  
 RL Blochum. Biophys. Acta 1437:333-340(1999).  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC IPP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE, AN IMPORTANT  
 CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE  
 CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE  
 CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY  
 CC -1- SUBUNIT: HOMOCYCLAMER  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC  
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN TESTIS. FOUND IN OTHER  
 CC TISSUES TO A LOWER EXTENT.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL: AB017971; BAA75909.1;  
 DR EMBL: AB019036; BAA77251.1;  
 DR EMBL: AF125394; AAD3030.1;  
 DR EMBL: AB016043; BAA6511.1;  
 DR InterPro: IPR000092;  
 DR Pfam: PF00348; PolyPrenyl\_Synthet\_1;  
 DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1;  
 DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2;  
 DR Isoprene biosynthesis; Transferase  
 KW SEQUENCE 300 AA; 34871 MW; FSD1959274BEE27A CRC64;  
 SQ

Query Match 100.0%; Score 1566; DB 1; Length 300;  
 Best local similarity 100.0%; Pred. No. 3e-111;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MKETQVORILPEYKYLQLPQKQVTRKLSQAFNMHLKVPEDKQIIIEVTMLHNAS 60  
 DB 1 MKETQVORILPEYKYLQLPQKQVTRKLSQAFNMHLKVPEDKQIIIEVTMLHNAS 60  
 OY 61 LIIDIEDNSKLRGPPVAHSIYIGIPSVINSANYVYFGLERKVLFLDHPDAVKLFTROLL 120  
 DB 61 LIIDIEDNSKLRGPPVAHSIYIGIPSVINSANYVYFGLERKVLFLDHPDAVKLFTROLL 120  
 OY 121 ELHOGGGLDIYWRNDYTCPTTEERYKAMVLOKTGGLFGLAVGIMOLFSDYKEDLKLNTL 180  
 DB 121 ELHOGGGLDIYWRNDYTCPTTEERYKAMVLOKTGGLFGLAVGIMOLFSDYKEDLKLNTL 180  
 OY 181 GLEFOIRDYANLHNSKEYSENKSFCEDLTEGKFSPTTHAHSRPSSTOVONILROREND 240  
 DB 181 GLEFOIRDYANLHNSKEYSENKSFCEDLTEGKFSPTTHAHSRPSSTOVONILROREND 240  
 OY 241 IDIKKCYVHLEVDVGSFEYTRNTLKELEKAKYKQIDARGNPELVALVYKLSKMKRENE 300  
 DB 241 IDIKKCYVHLEVDVGSFEYTRNTLKELEKAKYKQIDARGNPELVALVYKLSKMKRENE 300  
 RESULT 2  
 ID GGPB\_BOVIN STANDARD; PRT; 294 AA.  
 AC P56966;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPSSASE)  
 DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
 DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE  
 DE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29)] (FRAGMENT)  
 CN GGPST.  
 OS Bos taurus (Bovine). Chordata: Vertebrata: Euteleostomi;  
 OC Eukaryota: Metazoa: Cetartiodactyla, Ruminantia, Pecora: Bovidae;  
 OC Mammalia: Eutheria: Cetartiodactyla, Ruminantia, Pecora: Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RP [1]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-14 AND 287-294.  
 RX MEDLINE=99150380; PubMed=10026212;  
 RA Kuzuguchi T., Morita Y., Sagami T., Sagami H., Ogura K.;  
 RT "Human geranylgeranyl diphosphate synthase. cDNA cloning and  
 RT expression."  
 RL J. Biol. Chem. 274:5888-5894(1999).  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC IPP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE, AN IMPORTANT  
 CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE  
 CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE  
 CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY  
 CC -1- SUBUNIT: HOMOCYCLAMER (BY SIMILARITY)  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL)  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 CC Isoprene biosynthesis; Transferase.  
 KW NON\_TER 1  
 FT NON\_TER 294  
 SQ SEQUENCE 294 AA; 34138 MW; IAC128C113FE8D53 CRC64;

Query Match 96.8%; Score 1516; DB 1; Length 294;  
 Best local similarity 98.3%; Pred. No. 1.7e-107;  
 Matches 289; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

OY 3 KTOETVORILPEYKYLQLPQKQVTRKLSQAFNMHLKVPEDKQIIIEVTMLHNAS 62  
 DB 1 KTOETVORILPEYKYLQLPQKQVTRKLSQAFNMHLKVPEDKQIIIEVTMLHNAS 60  
 OY 63 IDIEDNSKLRGPPVAHSIYIGIPSVINSANYVYFGLERKVLFLDHPDAVKLFTROLL 122  
 DB 61 IDIEDNSKLRGPPVAHSIYIGIPSVINSANYVYFGLERKVLFLDHPDAVKLFTROLL 120  
 OY 123 HOGGGLDIYWRNDYTCPTTEERYKAMVLOKTGGLFGLAVGIMOLFSDYKEDLKLNTL 182  
 DB 121 HOGGGLDIYWRNDYTCPTTEERYKAMVLOKTGGLFGLAVGIMOLFSDYKEDLKLNTL 180  
 OY 183 FFOIRDYANLHNSKEYSENKSFCEDLTEGKFSPTTHAHSRPSSTOVONILROREND 242  
 DB 181 FFOIRDYANLHNSKEYSENKSFCEDLTEGKFSPTTHAHSRPSSTOVONILROREND 240  
 OY 243 IKKCYVHLEVDVGSFEYTRNTLKELEKAKYKQIDARGNPELVALVYKLSKMKR 296  
 DB 241 IKKCYVHLEVDVGSFEYTRNTLKELEKAKYKQIDARGNPELVALVYKLSKMKR 294

RESULT 3  
 ID GGPB\_MOUSE STANDARD; PRT; 300 AA.  
 AC Q9WTNO;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPSSASE)  
 DE (GERANYLGERANYL DIPHOSPHATE SYNTHASE) [INCLUDES:  
 DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE  
 DE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29)]



GN GCP51.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Lymph node;  
RX MEDLINE=99203156; PubMed=10101267;  
RA Kainou T., Kawamura K., Tanaka K., Matsuda H., Kawanukai M.;  
RT "Identification of the GCP51 genes encoding geranylgeranyl diphosphate  
synthases from mouse and human."  
RL Biochem. Biophys. Acta 1437:333-340(1999).  
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE. AN IMPORTANT  
CC PRECURSOR OF CAROTENOIDS AND GERANYLATEDGERANYLATED PROTEINS.  
CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
CC -1- PATHWAY: ISOPRENOID BIOSYNTHETIC PATHWAY.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AB016044; BAA76512.1;  
DR MOD: MGI:1341724; Gcp51.  
DR InterPro: IPR000092;  
DR Pfam: PF00348; Polyprenyl\_synth\_1;  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
KW Isoprene biosynthesis; Transferase.  
SQ SEQUENCE 300 AA; 34707 MW; 13ECB67EA17EE63 CRC64;

Query Match 94.3%; Score 1476; DB 1; Length 300;  
Best Local Similarity 93.7%; Pred. No. 1,8e-104;  
Matches 281; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

QY 1 MEKTOETVORILEPYYVLLQPGKOVRTKLSOAFNHLKVPEDKQIIEVTEMLHNAS 60  
DB 1 MEKTKAEKRLLEPYRYLLQPGKOVRSKLSOAFNHLKVPEDKQIIEVTEMLHNAS 60  
QY 61 LLIIDIDNSKLRGEPVASHYIGIPVINSANYVYELGLEKVLTLDPDAVKLFTROLL 120  
DB 61 LLIIDIDSSKLRGEPVASHYIGIPVINSANYVYELGLEKVLTLDPDAVKLFTROLL 120  
QY 121 ELHOGGGLDIYWRDNYTCPEEERKANYLQKGTGGLAVGLMQLFSDYKEDKPLINTL 180  
DB 121 ELHOGGGLDIYWRDNYTCPEEERKANYLQKGTGGLAVGLMQLFSDYKEDKPLINTL 180  
QY 181 GLFPOJDDYANLHRSKESYENKSCEDLTGKFSFPTIHAIWSPRESTOVONILKORTEN 240  
DB 181 GLFPOJDDYANLHRSKESYENKSCEDLTGKFSFPTIHAIWSPRESTOVONILKORTEN 240  
QY 241 IDIKRYCVYLEDVGSFAYTRHTLRELEAKAYKQIEACGNSPLVALVKHLSKMFTEENK 300  
DB 241 IDIKRYCVYLEDVGSFAYTRHTLRELEAKAYKQIEACGNSPLVALVKHLSKMFTEENK 300

RESULT 4  
GGP\_NEUCR STANDARD; PRT; 428 AA.

AC P24322;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GCP51) (GCP51)  
DE (GERANYLGERANYL DIPHOSPHATE SYNTHETASE) [INCLUDES:  
DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE  
DE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29)].  
GN AL-3 OR ALBINO-3.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
OX NCBI\_TaxID=5141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=74-OR23-1A;  
RX MEDLINE=91170267; PubMed=1826006;  
RA Caratelli A., Romano N., Ballario P., Morelli G., Macino G.;  
RT "The Neurospora crassa carotenoid biosynthetic gene (albino 3)  
RT reveals highly conserved regions among prenyltransferases.";  
RL J. Biol. Chem. 266:5854-5859(1991).  
CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
CC -1- PATHWAY: CAROTENOID BIOSYNTHESIS.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- INDUCTION: BY BLUE LIGHT.  
CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U20940; AAC13867.1;  
DR PIR: S15662; S15662.  
DR InterPro: IPR000092;  
DR Pfam: PF00348; Polyprenyl\_synth\_1;  
DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
KW Carotenoid biosynthesis; Isoprene biosynthesis; Transferase.  
FT ACT SITE 268 268 BY SIMILARITY.  
SQ SEQUENCE 428 AA; 47887 MW; 7989DEB8BDE360F CRC64;

Query Match 50.9%; Score 797.5; DB 1; Length 428;  
Best Local Similarity 50.7%; Pred. No. 4.5e-53;  
Matches 151; Conservative 57; Mismatches 85; Indels 5; Gaps 2;

QY 6 ETVORILEPYYVLLQPGKOVRTKLSOAFNHLKVPEDKQIIEVTEMLHNASLIDD 65  
DB 123 EEKEXVLTGVDYDNGHPRGDISQAVKAPDAWLDVSESELEVITKVISMLHTASLLYDD 182  
QY 66 IEENSKLRGEPVASHYIGIPVINSANYVYELGLEKVLTLDPDAVKLFTROLLHOG 125  
DB 183 VEDNSVILKRGEPVASHYIGIPVINSANYVYELGLEKVLTLDPDAVKLFTROLLHOG 242  
QY 126 QGLDIYWRDNYTCPEEERKANYLQKGTGGLAVGLMQLFSDYKEDKPLINTLGLFQ 185  
DB 243 QGMDLFWRDLTCTPEDDYLEMVSNKGTGGLFGLIKLMQESRSPVDCVPLVNTIGLIFQ 302  
QY 186 IRDDYANLHRSKESYENKSCEDLTGKFSFPTIHAIWSPRESTOVONILKORTENIDIK 245  
DB 303 IADYHNLNMRVYANGMGCEDLTGKFSFPTIHAIWSPRESTOVONILKORTENIDIK 362

QY 246 YCHYLEDVGSFEYTRNTKELKELAKAYKQI-----DARGNPELVALVYKLSMKPEKEN 299  
 DB 363 YAVAMESTGSEFYTRKYLKVLVDRAROMTIDIDGRCGSGIKHILDRIMLHOEEN 419

RESULT 5  
 GGGP\_GIBFU STANDARD; PRT; 418 AA.

AC 092236;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGP SYNTHETASE) (GGPSPASE)  
 DE DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1); GERANYLTRANSFERASE  
 DE (EC 2.5.1.10); FARNESYLTRANSFERASE (EC 2.5.1.29).  
 GN GGS OR GGPS.  
 OS Glibberella fujikuroi (Fusarium moniliforme).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Hypocreales; Nectriaceae; Glibberella.  
 OX NCBI\_TaxID=5127;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=M67;  
 RX MEDLINE=97374453; PubMed=9230902;  
 RA Mende K., Homann V., Tudzynski B.;  
 RT "The geranylgernanyl diphosphate synthase gene of Glibberella fujikuroi:  
 RT isolation and expression.";  
 RT Mol. Gen. Genet. 255:96-105(1997).  
 RL [1]  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC 1TP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE.  
 CC CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
 CC CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
 CC PATHWAY: BIOSYNTHESIS OF GIBBERELLINS.  
 CC SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGP SYNTHETASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X96943; CAA65644.1; -  
 CC InterPro: IPR000092; -  
 CC Pfam: PF00348; polyprenyl synt. 1.  
 CC PROSITE: PS00723; POLYPRENYL SYNTHET. 1;  
 CC PROSITE: PS00444; POLYPRENYL SYNTHET. 2.1;  
 CC Isoprene biosynthesis; Transferase.  
 CC FT ACT\_SITE 269 BY SIMILARITY.  
 CC SEQUENCE 418 AA; 46477 MW; BE2A0F1EA7D176C9 CRC64;

QY 9 ORLLPPEYKLLDLPKGOVETKLSOAFNMHLKVPEDKLOIIIEVTEMLHNASLIDIDIED 68  
 DB 127 ENYVRGPYIVISHPKDFPAQLIGAFNWLADVPITSLVITTVGMLHSSLLIDIDYOD 186

QY 69 NSKLRLGEPVAHSIYSPVINSANTYVLEGLKVLTLDPADVAKLFTROLLELHOGGL 128  
 DB 187 SSBLRGRGEPVAHNIYFVACTINGNATYFVALQELHKLNNPELITLFSDELVALHGGQGM 246

QY 129 DIWVRNYNCPTEEEKKAWLOKTKGLFGLAVGLM--QLFSYKEDLKLPLNLGLGFPQI 186

Query Match 48.2%; Score 754.5; DB 1; Length 418;  
 Best Local Similarity 49.8%; Pred. No. 7.6e-50;  
 Matches 145; Conservative 53; Mismatches 86; Indels 7; Gaps 3;

DB 247 DLFWCDLTCTPTEEDYLEMVGNTKGLFRLGIKLMAEANGSPDPCVPLNLIGLFIQI 306  
 QY 187 RDDXANLHSEYSEYKNSKFCEDLTREGKSPFTIARINSPSTOVONLRORTENIDIKKY 246  
 DB 307 RDDYMLNLSKSEYSHNKCWEDLTREGKSPFTIARINSPSTOVONLRORTENIDIKKY 366

QY 247 YCHYLEDVGSFEYTRNTKELKELAKAYK---QIDARGNPELVALVYKLSM 294  
 DB 367 YAVAMESTGSEFYTRKYLKVLVDRAROMTIDIDGRCGSGIKHILDRIMLHOEEN 419

RESULT 6  
 IDSA\_METUA STANDARD; PRT; 327 AA.

AC 058270;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE BIFUNCTIONAL SHORT CHAIN ISOPRENYL DIPHOSPHATE SYNTHASE [INCLUDES:  
 DE FARNESYL PYROPHOSPHATE SYNTHETASE (EC 2.5.1.1) (FPP SYNTHETASE)  
 DE (DIMETHYLLALYLTRANSFERASE); GERANYLTRANSFERASE (EC 2.5.1.10)].  
 GN IDSA OR MJ0860.  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.L.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii.";  
 RL Science 273:1058-1073(1996).  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.  
 CC -1- COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGP SYNTHETASES FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U67530; AAB98865.1; -  
 CC TIGR: MJ0860;  
 CC InterPro: IPR000092; -  
 CC Pfam: PF00348; polyprenyl synt. 1.  
 CC PROSITE: PS00444; POLYPRENYL SYNTHET. 2.1;  
 CC PROSITE: PS00723; POLYPRENYL SYNTHET. 1;  
 CC Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;  
 CC Multifunctional enzyme.  
 CC SEQUENCE 327 AA; 37568 MW; 07D68AC9BD657DAC CRC64;

QY 186 YCHYLEDVGSFEYTRNTKELKELAKAYKQI-----DARGNPELVALVYKLSMKPEKEN 299  
 DB 363 YAVAMESTGSEFYTRKYLKVLVDRAROMTIDIDGRCGSGIKHILDRIMLHOEEN 419

Query Match 18.6%; Score 292; DB 1; Length 327;  
 Best Local Similarity 31.0%; Pred. No. 4.3e-15;  
 Matches 85; Conservative 56; Mismatches 111; Indels 22; Gaps 10;

QY 17 KYLLQPGKQVRFKLSQAFNHMLKVPEDKLOIIT---EVTMLHNASLIDIEDNSKR 73  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 40 KHLFAGCKRIPLT-VVYIMLK--KDIIEVLPAAAVELHNTLHDDIMDNDER 96  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 74 RGPVAHSITGIPSVINSANYVFLGLEKVLTL-DHPDA---VKLFTROLLEHOGGLD 129  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 97 RGRVTVHVVYGEPMALLAGLLYAKAFEAVSRIKDNKKAHEVLKLSKACVECEGAMD 156  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 130 IYRDNATCTPEEYKAMVLOKGTGFLGVLGMOLEFSYDKE---LKLPLNTLGLFQ 185  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 157 MEENYNY--FTMEYEDIMIRKKGALLASVIGAVMADCNBEREALKEVAKRIGLTFQ 214  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 186 IRDYNALHSKESSEKSPCEDLTEGKFSPTTHAISRPE--STOVONTLRQNTENIDI 243  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 215 IODVDLIDGOKKLGKPGVSDIREGKTIIVHIALKTLDEDKKRLLETLGKNKYNKDEE 274  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 244 KYCVHYLEDVGSFEYTRNTLKELE--EAKAYKOI 275  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 275 IKRAEILKP--SIEYAKELMKOTEAKKEYLKI 306  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :

RESULT 7  
 IDSA\_METHM STANDARD; PRT; 324 AA.  
 AC 053479;  
 ID 01-NOV-1997 (Rel. 35, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE BIFUNCTIONAL SHORT CHAIN ISOPRENYL DIPHOSPHATE SYNTHASE [INCLUDES:  
 DE FARNESYL PYROPHOSPHATE SYNTHETASE (EC 2.5.1.1) (FPP SYNTHETASE)  
 DE (DIMETHYLLALYLTRANSFERASE); GERANYLTRANSFERASE (EC 2.5.1.10).]  
 GN IDSA.  
 OS Methanobacterium thermoautotrophicum (strain Matburg / DSM 2133).  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacter.  
 RX NCBI\_TaxID=79929;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=95070145; PubMed=7979381;  
 RA Chen A., Poulier C.D.;  
 RT "Isolation and characterization of *Idsa*: the gene for the short chain  
 RT isoprenyl diphosphate synthase from *Methanobacterium*  
 RT thermoautotrophicum";  
 RL Arch. Biochem. Biophys. 314:399-404(1994).  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-FARNESYL DIPHOSPHATE.  
 CC -1- COFACTOR: REQUIRES MAGNESIUM.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: S75695; AAB32421.1; -  
 CC DR InterPro: IPR000092; -  
 CC DR Pfam: PF000348; polyprenyl\_synth\_1;  
 CC DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1;  
 CC DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1;  
 CC DR Lipid synthesis: Isoprene biosynthesis; Transferase; Magnesium;  
 CC KM Multifunctional enzyme.  
 CC FT INIT MET 0  
 CC SEQUENCE 324 AA; 35505 MW; 88558A84DF5F6C1 CRC64;  
 SO

Query Match

14.9%; Score 233.5; DB 1; Length 324;

Best Local Similarity 27.5%; Pred. No. 1,1e-10;  
 Matches 84; Conservative 55; Mismatches 125; Indels 41; Gaps 13;  
 QY 1 MEKTOETVQRIILEPYKRYLLQPGKQVRFKLS---QAFNHMLKVPEDKLOIITEVTEML 56  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 19 MECISDITPDLTKASHEHLLTAGKRIKIPSLALISCBAV---GNPEDAGVAAAI-ELI 74  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 57 HNASLIDIEDNSKLRGPRVAHSITGIPSVINSANYVFLGLEKVLTLH---PDAY 112  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 75 HTFSLIHDDIMDDENKRGSPSVHIVGEPMAILAGVLFESKAEVAVIRNGDSERVDAL 134  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 113 KLFTROLLEHOGGLDIYRDNATCTPEEYKAMVLOKGTGFLGVL---GLMOLEFSY 169  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 135 AAVYDSCVKICEQALDMGEERLDV--TEDEYEMIKKRAALIAATKAGALMGASER 193  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 170 K-EDLAPLNTLGLFQIRDYNALHSKESSEKSPCEDLTEGKFSPTTHAISRPEST 228  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 194 EVELAEDYGRFGLAQIHDDYLDVVSDESLKPGVSDIREGKMTLMVYKALEASEED 253  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 229 QVONTLRQ-----TENIDI-KKY-CVHYLEDVGSFEYTRNTLKELEAKAY 273  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 254 -----RERLISILSGDEGSVAEALIEFERGATGYAHEV-ALDYVRMKEKLE--TLE 304  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 QY 274 QIDAR 278  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 DB 305 DSDAR 309  
 ||| ||| : : : : : ||| : : : : : ||| : : : : :  
 RESULT 8  
 IDSA\_METHM STANDARD; PRT; 325 AA.  
 AC 026156;  
 ID 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE BIFUNCTIONAL SHORT CHAIN ISOPRENYL DIPHOSPHATE SYNTHASE [INCLUDES:  
 DE FARNESYL PYROPHOSPHATE SYNTHETASE (EC 2.5.1.1) (FPP SYNTHETASE)  
 DE (DIMETHYLLALYLTRANSFERASE); GERANYLTRANSFERASE (EC 2.5.1.10).]  
 GN IDSA OR MTH50.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanobacter.  
 RX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DELTA H;  
 RX MEDLINE=98037514; PubMed=93711463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Nollan J., Reeve J.N.;  
 RT "Complete genome sequence of *Methanobacterium thermoautotrophicum*  
 RT deltaH: functional analysis and comparative genomics";  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE  
 CC = PYROPHOSPHATE + TRANS-FARNESYL DIPHOSPHATE.  
 CC -1- COFACTOR: REQUIRES MAGNESIUM.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AE000797; AAB84557.1;  
 DR InterPro; IPR000092;  
 DR Pfam; PF00348; Polyprenyl\_synth; 1.  
 DR PROSITE; PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 DR Lipid synthesis; Isoprene biosynthesis; Transferase; Magnesium;  
 KM Multifunctional enzyme.  
 SQ SEQUENCE 325 AA; 35485 MW; BCF3A285A207916E CRC64;

Query Match 14.8%; Score 231.5; DB 1; Length 325;  
 Best Local Similarity 25.0%; Pred. No. 1.6e-10;  
 Matches 79; Conservative 59; Mismatches 133; Indels 45; Gaps 10;

OY 2 EKQETVQRI---LLEYYKYLLOPGQVPTKLSQAFNNHLKPKEDKLTQIIIEVTEKHL 57  
 Db 17 ERRESISDITPEPILLRASEHLLITRAGKRRPSIALISSEAVGDDPDAGVAAAIELIH 76  
 OY 58 NASLLIDIEDNSKLRGFPVAHSITGIPSVINSANYVFLGLEKVLTDHPDAVK--- 113  
 Db 77 TFSLLHDIMDDDIRGEPVAHVHNGEPMAIIAGDVLFSKAFPAVIRNGDSEWKEPALA 136  
 OY 114 LFTROLLEHOGGLDIYWRDNYTCPTPEEEKKAMVLOKTGGLFGLAVGL----- 162  
 Db 137 VVVDSCVKICEGQALDMGFEBRLDV-TEEEYEMKIYKKTALIAAATKAGAIMGGSPQE 195  
 OY 163 MQLESDEKEDLKPLNTLGLFQIORDYANILHSKEYSENKSFCDLTGEGKFSPTTHAI- 221  
 Db 196 IAALEDG-----RCIGLAFIHDYDLYVSDSESLGKPVGSDIAGCKMTLMVYKALE 248  
 OY 222 -WSRPESTOVONILRORTENIDIKKCVHLEDDVGSPEY-----TRMTLELEKAY 272  
 Db 249 RASKDKBERLISILSGDEKIVAE-AIEIFERIGATEYAAVAALDHYRAKEKLE--VL 304  
 OY 273 KQIDARGNPELVALV 288  
 Db 305 EESDAR---EALAMI 316

RESULT 9  
 PREA\_CYAPA STANDARD; PRT; 323 AA.  
 ID PREA\_CYAPA  
 AC P31171;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE PRENYL TRANSFERASE (EC 2.5.1.-).  
 GN PREA.  
 OS Cyanophora paradoxa.  
 OG Cyanelle.  
 OC Eukaryota; Glaucocystophyceae; Cyanophoraceae; Cyanophora.  
 OX NCBI\_TaxID=2762;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / PRINGSHEIM;  
 RX MEDLINE=91268060; PubMed=1711042;  
 RA Michalowski C.B., Loeffelhardt W., Bohnert H.J.;  
 RT "An ORF323 with homology to crte, specifying prephytoene  
 pyrophosphate dehydrogenase, is encoded by cyanelle DNA in the  
 eukaryotic alga Cyanophora paradoxa."  
 RT J. Biol. Chem. 266:11866-11870(1991).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN-LB555 / PRINGSHEIM;  
 RA Stirewalt V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,  
 RL Bryant D.A.;  
 RT Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-46 FROM N.A.  
 RC STRAIN-UTEX 5550;  
 RX MEDLINE=91117189; PubMed=2126059;  
 RA Michalowski C.B., Pflanzagl B., Loeffelhardt W., Bohnert H.J.;

RT "The cyanelle S10 spe ribosomal protein gene operon from Cyanophora  
 paradoxa."  
 RL Mol. Gen. Genet. 224:222-231(1990).  
 CC -1- FUNCTION: POSSIBLE ROLE IN SYNTHESIS OF THE NONAPRENYL SIDE CHAIN  
 OF PLASTOCOINONE OR IN SYNTHESIS OF OTHER PRENYL CHAINS SUCH AS  
 UNDECAPEPTYL PYROPHOSPHATE  
 CC -1- SIMILARITY: BELONGS TO THE PPP/GGPP SYNTHETASES FAMILY.  
 CC -1- CAUTION: WAS ORIGINALLY (REF.1) CALLED CRTE.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M37111; AAA65472.1;  
 DR EMBL; U30821; AAA81217.1;  
 DR EMBL; M30487; AAA63631.1;  
 DR PIR; A40433; A40433.  
 DR InterPro; IPR000092;  
 DR Pfam; PF00348; Polyprenyl\_synth; 1.  
 DR PROSITE; PS00444; POLYPRENYL\_SYNTHET\_2; 1.  
 DR PROSITE; PS00723; POLYPRENYL\_SYNTHET\_1; 1.  
 DR Photosynthesis; Isoprene biosynthesis; Transferase; Cyanelle.  
 KW SEQUENCE 323 AA; 35919 MW; CB480246342809A CRC64;

Query Match 13.9%; Score 217; DB 1; Length 323;  
 Best Local Similarity 29.8%; Pred. No. 1.9e-09;  
 Matches 67; Conservative 36; Mismatches 108; Indels 14; Gaps 5;

OY 49 IEVTEMLNASSLLIDIEDNSKLRGFPVAHSITGIPSVINSANYVFLGLEKVLTDH 108  
 Db 73 LAETIEIHTASLVHDDIIDEQVRGIPVHSDEFTKAILAGDPLFQSSWYLANLES 132  
 OY 109 PDAVKLFTROLLEHOG---GLDIYWRDNYTCPTPEEEKKAMVLOKTGGLFGLAVGLMOL 165  
 Db 133 LEVYKLSIKSVITDFAGELIRGLNOKFVD---LLEEVLEKSFYKTSILAASSKAAL 188  
 OY 166 FS---DYKEDLKPLNTLGLFQIORDYANILHSKEYSENKSFCDLTGEGKFSPTTHAI 221  
 Db 189 LSHVDLTVANLDLYNGRHGLAFQIYVDILDPTSSIEELGKSCSDLKGNLTAPVLRAL 248  
 OY 222 WSRPESTOVONILRORTENIDIKKCVHLEDDVGSFEYTRNTLKE 266  
 Db 249 -EONSLEIPLIOROFSEPKDF-EYTLQIVEETKAIETKRELAME 290

RESULT 10  
 GGPP\_SULSO STANDARD; PRT; 332 AA.  
 ID GGPP\_SULSO  
 AC P95999;  
 DT 01-OCT-2000 (Rel. 40, Created)  
 DT 01-OCT-2000 (Rel. 40, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGPP SYNTHETASE)  
 DE [INCLUDES: DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1);  
 DE GERANYLTRANSFERASE (EC 2.5.1.10); FARNESYLTRANSFERASE  
 DE (EC 2.5.1.29)].  
 GN GDS OR C05010.  
 OS Sulfolobus solfataricus.  
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 OX NCBI\_TaxID=2287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-DSM 1617 / P2;  
 RX MEDLINE=97055432; PubMed=8899719;  
 RA Sensen C.W., Klenk H.-P., Singh R.K., Allard G., Chan C.C.-Y.,  
 RA Liu Q.Y., Penny S.L., Young F., Schenk M.E., Gaasterland T.,  
 RA Doolittle W.F., Ragan M.A., Charlebois R.L.;  
 RT "Organizational characteristics and information content of an archaeal

RT genome: 156 kb of sequence from *Sulfolobus solfataricus* p2.;

RL Mol. Microbiol. 22:175-191(1996).

CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF

CC ITP ONTO DMAP TO FORM GERANYLGERANYL PYROPHOSPHATE WHICH IS A

CC PRECURSOR OF THE ETHER-LINKED LIPIDS.

CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL

CC DIPHOSPHATE = PYROPHOSPHATE + GERANYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL DIPHOSPHATE

CC = PYROPHOSPHATE + TRANS-TRANS-FARNESYL DIPHOSPHATE.

CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL

CC DIPHOSPHATE = PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.

CC -1- PATHWAY: BIOSYNTHESIS OF MEMBRANE ETHER-LINKED LIPIDS.

CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: Y08257; CAA69541.1; -

CC InterPro: IPR000092; -

CC DR Pfam: PF00348; Polyprenyl\_synth; 1.

CC DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.

CC DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.

CC KW Lipid synthesis; Isoprene biosynthesis; Transferase;

CC MW Multifunctional enzyme.

CC SQ SEQUENCE 332 AA; 37259 MW; ADA93E594E014A60 CRC64;

Query Match 13.4%; Score 209.5; DB 1; Length 332;

Best Local Similarity 26.6%; Pred. No. 7.4e-09;

Matches 75; Conservative 53; Mismatches 125; Indels 29; Gaps 9;

QY 10 RILEPEKYLLPGKQVTRKLSQAFNMVKVPEDKLQITTEVEMAHNLSLIDIEDN 69

DB 30 KTLSEASFHLFTAGKRLRLVLVSSDLGCDREKAYKAAVAELIANTLVHDDMDN 89

QY 70 SKLRGFPVAHSYIGIPSVINSANYV---FLGL-EKVLTD---HPDAVKLFTRQLLEL 122

DB 90 DGLRGLPTVHVKMGEMALTAGDYTLHAKAFESINSLKGDGNTFYKARSIFTSIEII 149

QY 123 HQOGGLDIYRNDYTCETEERYKAMVLOKGTGLAVGLMOLFSDYKEDLKLPLN--- 178

DB 150 SEQQAMDMSE-ENRDLVTEERYIOMIKGTAMLFSCSAALGGIINKANDVYKRLTEYGL 208

QY 179 TLGLFQIRDDYANLHSEKSENNKSFCEDETEGKFSPTTHAIMSRPESTOVONIL--- 234

DB 209 NLGSPQIVDDIIGIDDELKGLKPIYSDIREGK---KTLIVIKITLSEATEDEKILVST 265

QY 235 ----RQTEINID----IKRYCVHYLEDVGSFEYTRNTLKEI 267

DB 266 LGNKEAKKEDLERASEIIRKHSIQAYADLAK-KYSDLAIEINL 306

RESULT 11

ISPB\_ECOLI

ID ISPB\_ECOLI STANDARD; PRT; 323 AA.

AC P19641;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1994 (Rel. 28, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE OCTAPRENYL-DIPHOSPHATE SYNTHASE (EC 2.5.1.-) (OCTAPRENYL PYROPHOSPHATE

DE SYNTHETASE) (OPP SYNTHETASE).

OS ISPB OR CEL.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN-K12 / W3110;

RX MEDLINE=94146411; PubMed=8312607;

RA Jeong J.H., Kitakawa M.S., Isono S., Isono K.;

RT "Cloning and nucleotide sequencing of the genes, rpiU and rpnA, for

RT ribosomal proteins L21 and L27 of *Escherichia coli*."

RL DNA Seq. 4:59-67(1993).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;

RT "The complete genome sequence of *Escherichia coli* K-12.";

RL Science 277:1453-1474(1997).

RN [3]

RP SEQUENCE OF 196-323 FROM N.A.

RX MEDLINE=89359178; PubMed=2670911;

RA Choi Y.-L., Nishida T., Kawamukai M., Utsumi R., Sakai H., Komano T.;

RT "Cloning and sequencing of an *Escherichia coli* gene, nlp, highly

RT homologous to the ner genes of bacteriophages Wu and D108.";

RL J. Bacteriol. 171:5222-5225(1989).

RN [4]

RP FUNCTION.

RC STRAIN-K12 / JM109;

RX MEDLINE=94311902; PubMed=8037730;

RA Asai K.-I., Fujisaki S., Nishimura Y., Nishino T., Okada K.,

RA Nakagawa T., Kawamukai M., Matsuda H.;

RT "The identification of *Escherichia coli* ispb (cel) gene encoding the

RT octaprenyl diphosphate synthase.";

RL Biochem. Biophys. Res. Commun. 202:340-345(1994).

CC -1- FUNCTION: SUPPLIES OCTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE

CC SIDE CHAIN OF THE ISOPRENOID QUINONES UBIQUINONE AND MENAQUINONE.

CC -1- SIMILARITY: BELONGS TO THE FPP/GGPP SYNTHETASES FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL: D13267; -; NOT ANNOTATED\_CDS.

CC EMBL: U18997; AAA57988.1; -

CC DR EMBL: AB000399; AAC76219.1; -

CC DR EMBL: X68873; CAA48735.1; ALT\_SEQ.

CC PIR: PV0010; PV0010.

CC EcoGene: EG10017; ispb.

CC InterPro: IPR000092; -

CC DR Pfam: PF00348; Polyprenyl\_synth; 1.

CC DR PROSITE: PS00444; POLYPRENYL\_SYNTHET\_2; 1.

CC DR PROSITE: PS00723; POLYPRENYL\_SYNTHET\_1; 1.

CC KW Isoprene biosynthesis; Transferase.

CC SQ SEQUENCE 323 AA; 35217 MW; 08AD7AE7AC230EF8 CRC64;

Query Match 13.3%; Score 208.5; DB 1; Length 323;

Best Local Similarity 24.8%; Pred. No. 8.5e-09;

Matches 75; Conservative 60; Mismatches 142; Indels 25; Gaps 8;

QY 1 MEKTOETVORILEPYKYLLQLPKQVTRKLSQAFNMVKVPEDKLQITTEVEMAHNLS 60

DB 22 LEQINSVQVLINDLGY-YVSGGKRRIRPMIAVLAARAVGEGNANHTVIALIEFIHTAF 80

QY 61 LLLDDIEDNSKLRGFPVAHSYIGIPSVINSANYVFLGLEKVLITDHPAVVLFTRQLL 120

DB 81 LLLDDVVDSEDMRGKATANAAGNAAVSLVGFYITRAFOQMTSLGSLKVLLEVMSEAVN 140

QY 121 ELHQOGGLDIYRNDYTCP---TEEYKAMVLOKGTGLAVGLMOLFSDYKEDLKLPLN 174

DB 141 VIAGEVYQLM---NVNDDPITEENYRWRYISKTARLFEAAACSGILACTPEEENGLO 197

OY 175 PLNTLTGLFQIRDYANLHSEKSEKNSFCEDLIEGKFSPTTHA-WSRPESTOVONT 233  
 DB 198 DVGRIYGLTAFQILDLDLNADGEOLGNKVGDLNEGKPTPLPLHAMHNGPEOAO---M 254  
 OY 234 LRORENTIDIKKYCVHYLEDV-----GSFEYTRNTLKELEKAKYKOIDARGNPDLVA 286  
 DB 255 IRTALEOGNGR-----HLEPVELEANNAGCSLEMTQORAEEDKAIALQVLPDTPMREA 310  
 OY 287 LV 288  
 DB 311 LT 312

RESULT 12  
 GGP\_SULAC STANDARD: PRT: 330 AA.  
 AC P39464.  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 01-OCT-2000 (Rel. 40, Last annotation update)  
 DE GERANYLGERANYL PYROPHOSPHATE SYNTHETASE (GGP SYNTHETASE)  
 DE [INCLUDES: DIMETHYLLALYLTRANSFERASE (EC 2.5.1.1);  
 DE GERANYLTRANSTRANSFERASE (EC 2.5.1.10); FARNESYLTRANSTRANSFERASE  
 DE (EC 2.5.1.29)].  
 GN GDS.  
 OS Sulfolobus acidocaldarius.  
 OC Archaea; Crenarchaeota; Sulfolobales; Sulfolobaceae; Sulfolobus.  
 NCBI\_TaxID=2285;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33909 / NCIB 11770 / DSM 639;  
 RA MEDLINE=94237902; PubMed=8182085;  
 RX Ohnuma S.-I., Suzuki M., Nishino T.  
 RT "Archaeobacterial ether-linked lipid biosynthetic gene. Expression  
 RT cloning, sequencing, and characterization of  
 RT geranylgeranyl-diphosphate synthase."  
 RT J. Biol. Chem. 269:14792-14797(1994).  
 CC -1- FUNCTION: CATALYZES THE TRANS-ADDITION OF THE THREE MOLECULES OF  
 CC IPP ONTO DMAPP TO FORM GERANYLGERANYL PYROPHOSPHATE WHICH IS A  
 CC PRECURSOR OF THE ETHER-LINKED LIPIDS.  
 CC -1- CATALYTIC ACTIVITY: DIMETHYLLALYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE - PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: GERANYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE - PYROPHOSPHATE + GERANYL DIPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: TRANS-TRANS-FARNESYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE - PYROPHOSPHATE + GERANYLGERANYL DIPHOSPHATE.  
 CC -1- PATHWAY: BIOSYNTHESIS OF MEMBRANE ETHER-LINKED LIPIDS.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGP SYNTHETASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D28746; BAA3200.1; -  
 CC PTR: A54058; A54058.  
 CC InterPro: IPR000092; -  
 CC DR Pfam: PF00348; POLYPRENYL\_Synth\_1;  
 CC DR PROSITE: PS00444; POLYPRENYL\_Synthet\_2; 1.  
 CC DR PROSITE: PS00723; POLYPRENYL\_Synthet\_1; 1.  
 CC KW Lipid synthesis; Isoprene biosynthesis; Transferase;  
 CC Multifunctional enzyme.  
 CC SEQUENCE 330 AA; 36874 MW; 69EEDACDABALBCB CRC64;

Query Match 13.2%; Score 207; DB 1; Length 330;  
 Best Local Similarity 24.7%; Pred. No. 1.1e-08;  
 Matches 74; Conservative 52; Mismatches 131; Indels 42; Gaps 10;

OY 12 LLEPKYLLQDPGQVTRTKLSQAFENHMLKVPEDKIQIIEVTEMLHNASLLIDIDEDNSK 71  
 DB 30 LYEASVHLFTSGKRLRPLILFTISSDLFGGGRERAYAGAIAEVLHFTLVHDDIMQDN 89  
 OY 72 LRGEFPAVASTYIGPSYINSANYV---FLGLEKVL-LDHPDAVK--LFTQOLLEHQ 124  
 DB 90 IRRGLPVTYHYVYGLPLAILMGDDLHAKAFOLLTQALGLPSEFTIRKAFDIFRTSIILIE 149  
 OY 125 GQGLDIWRDNYTCPTTEEEKAWYLQKGGFLGAVL-----NOLSPDKEDL 173  
 DB 150 GQAVDMFEEDRID-KQOEVLDMISRTALFSASSSIGALLGANDNDVRLMSDFG--- 205  
 OY 174 KPLNTLTGLFQIRDYANLHSEKSEKNSFCEDLIEGKFSPTTHA--WSRPESTOVQ 231  
 DB 206 ---TNGIAFOYVDLILGLTADEKELGKVFESDIRGKRTIIVIKYLECKDEKKIYL 261  
 OY 232 NLRORFENID-----IKKYCVHYLEDVGSFEYTRNTLKELEKAKYKOIDARGNP 282  
 DB 262 KALGNKSAKSEELMSSADITIKTSLDYAVLAE-KYKNAIDSL-----NOVSSKSDIP 314

RESULT 13  
 HEP2\_BACSU STANDARD: PRT: 348 AA.  
 ID HEP2\_BACSU  
 AC P31114.  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE HEPTRAPRENYL DIPHOSPHATE SYNTHASE COMPONENT II (EC 2.5.1.30) (HEPPP  
 DE SYNTHASE SUBUNIT 2) (SPORE GERMINATION PROTEIN C3).  
 GN HEP2 OR HEP3 OR GERC3.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Hennner D.J.;  
 RT "Sequence of Bacillus subtilis dbpA, mtr(A,B), gerc(1-3), ndk, cher,  
 RT aro(B,E,F,H), trp(A-F), hish, and tyrA genes."  
 RT submitted (JAN-1992) to the EMBL/genbank/DBJ databases.  
 RN [2]  
 RP CHARACTERIZATION OF GERC LOCUS  
 RX MEDLINE=91037938; PubMed=2121900;  
 RA Yazdi M.A., Moir A.;  
 RT "Characterization and cloning of the gerc locus of Bacillus subtilis  
 RT 168."  
 RT J. Gen. Microbiol. 136:1335-1342(1990).  
 RN [3]  
 RP FUNCTION.  
 RX MEDLINE=98386502; PubMed=9720033;  
 RA Leathersbarrow A.V.H., Yazdi M.A., Curson J.P., Moir A.;  
 RT "The gerc locus of Bacillus subtilis, required for menaquinone  
 RT biosynthesis, is concerned only indirectly with spore germination."  
 RT Microbiology 144:2125-2130(1998).  
 CC -1- FUNCTION: SUPPLIES HEPTRAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE  
 CC SIDE CHAIN OF THE ISOPRENOID OUTLINE MENAQUINONE-7 (MQ-7).  
 CC -1- CATALYTIC ACTIVITY: ALL-TRANS-HEXAPRENYL DIPHOSPHATE + ISOPENTENYL  
 CC DIPHOSPHATE - DIPHOSPHATE + ALL-TRANS-HEPAPRENYL DIPHOSPHATE.  
 CC -1- SUBUNIT: HETERODIMER OF COMPONENT I AND II.  
 CC -1- SIMILARITY: BELONGS TO THE FPP/GGP SYNTHETASES FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M80245; AAA20856.1; -  
 CC EMBL: Z99115; CAB14190.1; -

RESULT	14
ISPB_HAEIN	
ID	ISPB_HAEIN STANDARD; PRT; 329 AA.
AC	P44916;
DT	01-NOV-1995 (Rel. 32, Created)
DT	01-NOV-1995 (Rel. 32, Last sequence update)
DT	15-JUL-1999 (Rel. 38, Last annotation update)
DE	OCTAPRENYL-DIPHOSPHATE SYNTHASE (EC 2.5.1.-) (OCTAPRENYL PYROPHOSPHATE SYNTHETASE) (OPP SYNTHETASE).
GN	ISPB OR H10881.
OS	Haemophilus influenzae.
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC	Haemophilus.
OX	NCBI_TaxID=727;
RN	[1]
RP	SEQUENCE FROM N.A.
RP	STRAIN=RD / KW20 / ATCC 51907;
RC	MEDLINE=95350630; PubMed=7542800;
RC	Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C., FINE L.D., Fitchman J.L., Fuhmann J.L., Georgagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;
RT	Whole-genome random sequencing and assembly of Haemophilus
RT	influenzae Rd.;
RT	Science 269:496-512(1995).
CC	-1- FUNCTION: SUPPLIES OCTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
CC	SIDE CHAIN OF THE ISOPRENOID QUINONES UBIQUINONE AND MENAQUINONE
CC	(BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE FPP/GAPP SYNTHETASES FAMILY.

RESULT	15
HEP2_BACST	
ID	HEP2_BACST
AC	STANDARD;
AC	P53785;
DT	01-NOV-1997 (Rel. 35, Created)
DT	01-NOV-1997 (Rel. 35, Last sequence update)
DT	30-MAY-2000 (Rel. 39, Last annotation update)
DE	HEPTARENYL DIPHOSPHATE SYNTHASE COMPONENT II (EC 2.5.1.30) (HEPP
DE	SYNTHASE SUBUNIT 2).
GN	HEPT OR HEPS-2.
OS	Bacillus stearothermophilus.
OC	Bacteria; Firmicutes; Bacillus/Clostridium group;
OC	Bacillus/Staphylococcus group; Bacillus.
OX	NCBI_TaxID=1422;
RN	[1]
RP	SEQUENCE FROM N.A.
RP	STRAIN=ATCC 10149;
RX	MEDLINE=95355463; Pubmed=7629164;
RA	Kojke-Takeshita A., Koyama T., Obata S., Ogura K.:
RT	"Molecular cloning and nucleotide sequences of the genes for two
RT	essential proteins constituting a novel enzyme system for heptaprenyl
RT	diphosphate synthesis.";
RL	J. Biol. Chem. 270:18396-18400(1995).
CC	-1- FUNCTION: SUPPLEIES HEPTAPRENYL DIPHOSPHATE, THE PRECURSOR FOR THE
CC	CATALYTIC ACTIVITY: ALL-TRANS-HEXAPRENYL DIPHOSPHATE + ISOPENTENYL
CC	DIPHOSPHATE = DIPHOSPHATE + ALL-TRANS-HEPTAPRENYL DIPHOSPHATE.
CC	-1- SUBUNIT: HETERODIMER OF COMPONENT I AND II.
CC	-1- SIMILARITY: BELONGS TO THE FPP/GAPP SYNTHETASES FAMILY.

